

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
16 octobre 2003 (16.10.2003)

PCT

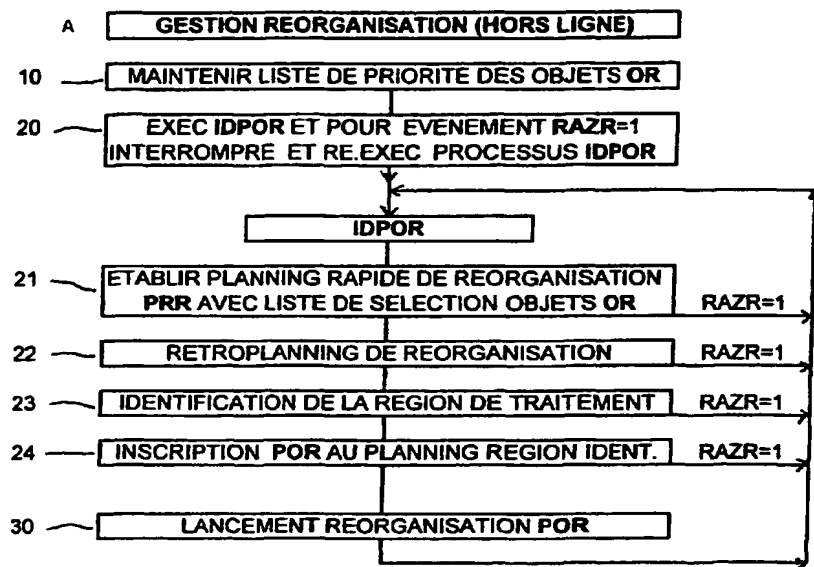
(10) Numéro de publication internationale
WO 03/085554 A1

- (51) Classification internationale des brevets⁷ : G06F 17/30 (71) Déposant (pour tous les États désignés sauf US) : INFO-
TEL [FR/FR]; -, 36 avenue du Général de Gaulle, Tour
(21) Numéro de la demande internationale : Gallieni II, F-93175 BAGNOLET (FR).
PCT/FR03/01076
(22) Date de dépôt international : 4 avril 2003 (04.04.2003) (72) Inventeur; et
(75) Inventeur/Déposant (pour US seulement) : LAFFORET,
(25) Langue de dépôt : français Bernard [FR/FR]; -, 13 avenue Alphonsine, F-94420 LE
PLESSIS TREVISE (FR).
(26) Langue de publication : français (74) Mandataire : CABINET JP COLAS; -, 37 avenue
Franklin D. Roosevelt, F-75008 PARIS (FR).
(30) Données relatives à la priorité : (81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ,
02/04301 8 avril 2002 (08.04.2002) FR BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ,

[Suite sur la page suivante]

(54) Title: METHOD FOR REORGANIZATION MANAGEMENT IN A SET OF INDEXED DATABASES OF A COMPUTER SYSTEM

(54) Titre : METHODE DE GESTION DES REORGANISATIONS DANS UN ENSEMBLE DE BASES DE DONNEES INDEXEES D'UN SYSTEME INFORMATIQUE



A...(OFFLINE) REORGANIZATION MANAGEMENT
10...MAINTAINING A PRIORITY LIST OF OBJECTS (OR)
20...EXECUTING (IDPOR) AND FOR EVENT RAZ=1
INTERRUPTING AND RE-EXECUTION OF PROCESS (IDPOR)
21...ESTABLISHING RAPID PLANNING OF
REORGANIZATION (PRR) WITH LIST OF OBJECTS (OR)
22...REORGANIZATION RETROACTIVE PLANNING
23...IDENTIFYING THE PROCESSING ZONE
24...WRITING (POR) IN IDENTIFIED PLANNING ZONE
30...LAUNCHING REORGANIZATION (POR)

(57) Abstract: The invention concerns an offline reorganization management in a set of indexed databases, in particular of the tablespace type, of an information computer system comprising the following steps: (10): continuously maintaining a list (PRIOREORG) of objects to be reorganized in descending order of priority; (20) executing a process (IDPOR) of identification of the next object to be reorganized (POR) in the absence of any reappraisal event (RAZ) and comprising the following steps: (21) establishing a rapid reorganization planning (PRR) using the list (PRIOREORG) and residual operational time (TRR) and a selection list (SELECT) of objects (OR); (22) establishing reorganization retroactive planning with LIFO memory storage of objects extracted from the list (SELECT); (23) identifying the available processing zone (RTR) to reorganize the object (POR); (24) writing (POR) when the zone (RTR) is cleared in the absence of any event (RAZ); and re-executing the process (IDPOR).

[Suite sur la page suivante]



DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LX, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) **États désignés-(régional)** : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK,

TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

- avec rapport de recherche internationale
- avant l'expiration du délai prévu pour la modification des revendications, sera republiée si des modifications sont reçues

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(57) **Abrégé** : La méthode de gestion des réorganisations " hors ligne " dans un ensemble de bases de données indexées, notamment de type espace de tables, d'un système informatique d'information comporte les opérations suivantes : (10) - Maintien en continu d'une liste PRIOREORG des objets à réorganiser par priorité décroissante; (20) - Exécution du processus IDPOR d'identification du prochain objet à réorganiser POR en l'absence d'événement de remise en cause RAZ et comportant les étapes suivantes : (21) - Etablissement d'un planning rapide de réorganisation PRR à l'aide de PRIOREORG et du temps opérationnel résiduel TRR et d'une liste de sélection SELECT d'objets OR; (22) - Etablissement d'un rétroplanning de réorganisation avec placement en mémoire LIFO des objets extraits de la liste SELECT ; (23) - Identification de la région de traitement RTR disponible pour réorganiser l'objet POR ; (24) - Inscription de POR au planning de la région identifiée ; (30) - Lancement réorganisation de POR dès libération de la région RTR en l'absence de tout événement RAZ; et re-exécution du processus IDPOR.

METHODE DE GESTION DES REORGANISATIONS DANS
UN ENSEMBLE DE BASES DE DONNEES INDEXEES
D'UN SYSTEME INFORMATIQUE

La présente invention concerne la réorganisation de bases de données gérées à l'aide d'un système informatique, plus particulièrement les bases de données indexées dans lesquelles un index, le plus souvent sous forme de table d'index, est utilisé pour accéder à un objet recherché, par exemple une ligne de la base de données. Ces bases de données peuvent être de très grande taille. Pour fixer les idées et sans caractère limitatif pour la présentation de l'invention la taille de ces bases de données de très grandes tailles peut aller de un million à quelques milliards de lignes. Ces bases de données sont en général enregistrées dans des mémoires de grande capacité, par exemple des mémoires à disque, chaque ligne de données étant adressée de façon appropriée. Les tailles considérables que peuvent avoir les bases de données requièrent souvent l'exploitation de ces dernières par partitions auxquelles sont associés des index de partitions spécifiques et/ou les index de bases de données complètes.

Il est capital, pour maintenir la capacité d'une base de données à être consultée par une connexion en ligne (on line) avec un temps d'accès aux données raisonnable, malgré l'effet retardateur du aux mises à jour pratiquées sur la base, de procéder à des réorganisations périodiques de cette base de données.

La réorganisation est une étape essentielle dans l'entretien des bases de données (et de leurs index). La réorganisation est faite hors ligne (off line) en général pendant la fenêtre de maintenance (fenêtre « batch ») et consiste à décharger la base de données, à ranger ses lignes d'information dans un ordre souhaité, par exemple l'ordre alphabétique, puis à recharger les lignes une fois rangées dans la base de données en mettant à jour les index, si possible de façon réorganisée.

Pour la suite de l'exposé on appellera indistinctement « objet à réorganiser » OR un objet susceptible de réorganisation d'un système d'information, en particulier la base de donnée seule, son ou ses index, l'association base de données et index, les partitions de la bases de données, leurs index associés (spécifiques ou non) et les associations partitions et leurs index.

De même pour la suite de l'exposé, on appellera système informatique d'information ou système d'information un système informatique complet (matériels et logiciels) adapté pour la gestion de bases de données et comportant au moins une base de données. On rencontre de tels systèmes notamment comme serveurs en ligne (on line) internes ou externes auxquels des utilisateurs se connectent pour consulter les bases de données correspondantes par l'intermédiaire de requêtes et autres demandes de recherche de données.

Les capacités « machine » du système d'information (unités centrales éventuellement mises en réseau) sont en général divisées en régions de traitement distinctes pilotées de façon unitaire par une région de contrôle. A chaque région de traitement sont affectées un certain nombre de tâches ordonnées (programmes ou sous-programmes à exécuter) et sont allouées les ressources matérielles nécessaires en mémoires virtuelles et périphériques (disques durs, imprimantes, mémoires sur supports externes etc...). Par souci d'efficacité en exécution et en allocation de ressources matérielles et/ou logicielles on affecte si possible à chaque région de traitement des tâches de même type, telles que par exemple pour des tâches généralement exécutables dans la fenêtre de maintenance dite fenêtre « batch », des réorganisations de fichiers ou bases de données (demandant beaucoup d'espace de disque) ou des copies images de fichiers (demandant des périphériques de sortie externe, lecteurs de bandes magnétiques, lecteurs/graveurs de CD-ROM, etc...).

D'une façon générale les réorganisations de bases de données sont nécessaires pour éviter que les performances des transactions (lecture et/ou mise à jour) accédant aux bases de données ne se dégradent. Toutefois il n'est pas possible de maintenir tous les objets d'un système d'information en état de parfaite organisation parce que :

- i) les réorganisations consomment du temps et des ressources du système (processeurs, mémoire vive et mémoires sur disques);
- ii) des réorganisations trop fréquentes entraîneraient des coûts de réalisations supérieurs aux surcoûts générés par l'exploitation de bases de données faiblement désorganisées.

Une première étape du travail de gestion de la réorganisation d'un ensemble de bases de données consiste en l'élaboration d'un planning d'exécution des réorganisations. Ce planning est utilisé pour le lancement des réorganisations des bases de données hors ligne (off line) pendant la période de maintenance journalière de quelques heures prévue pour le système d'information (fenêtre « batch » ou fenêtre des exécutions groupées).

On connaît des outils logiciels susceptibles de générer le planning d'exécution des réorganisations à venir en fonction :

- de la liste des objets à réorganiser,
- de la fréquence prédéterminée de réorganisation par objet,
- du niveau de désorganisation des objets mesuré par des techniques non documentées.

Toutefois aucun des outils connus ne remet en cause le planning initial en fonction d'évènements pouvant survenir pendant l'exécution des réorganisations planifiées.

Accessoirement aucun de ces outils ne prétend procéder à une étude d'optimisation pour générer ce planning, notamment à partir des ressources matérielles et logicielles disponibles instantanément et/ou à partir des coûts de réorganisation et/ou de non-réorganisation ou en fonction de paramètres instantanés réels ou estimés des objets à réorganiser tels que leur taille, leur taux de désorganisation et leur temps de réorganisation.

L'invention a pour but une gestion de la réorganisation d'un ensemble de bases de données indexées d'un système informatique d'information permettant d'améliorer les performances d'une configuration donnée du système d'information par une recherche d'optimisation continue, notamment du point de vue coûts et/ou des temps de réorganisation des réorganisations des bases de données (et/ou leurs partitions) et de leurs index.

A cette fin, l'invention propose une méthode de gestion des réorganisations dans un ensemble de bases de données indexées d'un système informatique d'information adaptée à la réorganisation « hors ligne » dans au moins une région de traitement de réorganisation du système et comportant les phases opérationnelles suivantes (voir figure 1) :

(10) - Création et maintien par mise à jour en continu d'une liste **PRIOREORG** des objets à réorganiser par priorité décroissante fonction de l'état de désorganisation des objets à réorganiser ;

(20) - Exécution du processus **IDPOR** d'identification du premier ou prochain objet à réorganiser **POR** avec, à chaque événement **RAZR=1** interne au système pour lequel la sélection du premier ou prochain objet à réorganiser doit être remise en cause, interruption et remise à zéro et nouvelle exécution du processus **IDPOR**, lequel processus comportant les étapes suivantes :

(21) - Etablissement d'un planning rapide de réorganisation **PRR** à l'aide de la liste **PRIOREORG** et du temps opérationnel résiduel disponible **TRR** dans l'ensemble des régions de réorganisation et d'une liste de sélection **SELECT/OR** d'objets à réorganiser **OR** ordonnés par priorités décroissantes optimisées en fonction du gain de la réorganisation de chaque objet **OR**, ledit gain étant défini comme le produit d'un facteur représentatif du taux de désorganisation d'un objet **OR** par le temps de réorganisation de cet objet;

(22) - Etablissement d'un rétroplanning de réorganisation dans lequel pour tout objet courant à réorganiser **OR** extrait de la liste **SELECT/OR** par priorité croissante pour favoriser le traitement avancé des objets de plus grandes tailles possibles, on calcule pour chaque région de réorganisation le dernier délai de la région **DDRR**, représentant dans la fenêtre « batch » allouée à la réorganisation « hors ligne » le

temps minimal nécessaire à la réorganisation de l'objet courant **OR**, égal dans le cas présent à la durée de temps consommé **DCRR** plus le temps de réorganisation de l'objet courant **DROR**;

5 (23) - Identification de la région de traitement **RTR** de l'objet courant **OR** par adéquation optimisée entre la durée de traitement disponible dans ladite région **DRR** avec la durée de traitement **DROR** nécessaire pour la réorganisation de l'objet **OR** avec **DRR - DRRR minimal ≤ 0** ;

10 (24) - Inscription de l'objet courant **OR** comme prochain objet à réorganiser **POR** au planning de la région identifiée en valorisant le pointeur **POR** correspondant à l'adresse de l'objet courant à réorganiser et en augmentant la durée de temps consommé **DCRR** dans la région du temps de la réorganisation **DROR**.

15 (30) - Lancement de la réorganisation de l'objet **POR** dès que la région de traitement de réorganisation **RTR** identifiée est libérée en l'absence de tout évènement **RAZR=1** survenu entre temps ; Remise à zéro et nouvelle exécution du processus **IDPOR**.

20 Ainsi l'introduction dans la méthode de gestion selon l'invention du principe même du planning rapide de réorganisation **PRR** (dont le temps machine est minimal pour permettre des réactualisations fréquentes sans véritable pénalisation) et du principe de la réactualisation permanente du planning rapide de réorganisation **PRR** permet de prendre en compte immédiatement l'évolution du système d'information complet et des ses principaux paramètres (taux instantané de désorganisations des objets à réorganiser, allongement de la durée d'une réorganisation par rapport à la durée estimée, libération d'une région de traitement, fin d'une tâche de réorganisation, etc) pour modifier les priorités des réorganisations et optimiser le lancement de celles-ci.

25 A côté des réorganisations de fichiers tels que bases de données les copies des fichiers modifiés constitue un impératif de sécurité dans le domaine de l'exploitation de systèmes informatiques. En effet pour des raisons de sécurité informatique certains systèmes d'information utilisés pour la gestion d'un ensemble de bases de données incorporent dans leur ensemble de logiciels d'exploitation un dispositif de verrouillage de la remise en service d'une base de données réorganisée avant la prise de la copie de sauvegarde correspondante.

30 Selon un premier mode d'exécution de la méthode de gestion selon l'invention l'ordre optimal des réorganisations est coordonné avec l'ordre des copies de sauvegarde d'objets du système d'information en tenant compte des réorganisations en cours ou juste effectuées. Selon une variante avantageuse l'ordre des copies de sauvegarde est modifié par l'exécution en première priorité, pour au moins un objet à

réorganiser **OR**, de la copie dudit objet **OR** au plus tôt en fin de traitement de réorganisation.

Pour ce faire, la méthode de gestion des réorganisations selon l'invention est globalisée pour gérer d'une part les réorganisations, d'autre part les copies sachant
5 que pour des raisons d'optimisation on a la faculté de découper les objets à traiter en partitions jusqu'aux plus petites parties d'objets à traiter possibles pour chaque type de traitement. Il en résulte que la liste et les objets à copier peuvent être différents de la liste et des objets à réorganiser.

Par ailleurs en ce qui concerne les copies, on distingue généralement trois
10 types de copie, la copie en continu ou copie « journal » (en anglais « log »), la copie incrémentale qui ne prend que les modifications apportées aux fichiers et la copie totale ou copie image. Dans le cadre de la présente invention, sont concernées principalement les copies images et accessoirement les copies incrémentales.

D'une façon similaire à ce qui a été mis en place pour les réorganisations, la
15 méthode de gestion des réorganisations et des copies comporte une phase de création et de maintien d'une liste **PRIOCOPIE** de priorité des objets à copier (en fonction du temps passé depuis la dernière copie de chaque objet à copier **OC**) et l'établissement d'un planning rapide de copie, sachant que certains objets **OR** juste réorganisés (par exemple une base de données) sont déverrouillés avant remise en ligne par une prise
20 de copie image; pour ce faire, on choisit dans le cadre de la présente invention de copier ces objets en **première priorité**. Il en résulte une double interaction entre le planning rapide de copie **PPC** et le planning rapide de réorganisation **PRR*** (voisin du planning **PRR** mais modifié en conséquence pour intégrer la composante copie), pour le premier, **PPC**, par la priorité première donnée à certaines ou à toutes les copies
25 après réorganisations et pour le second, **PRR***, par l'allongement du temps total de réorganisation par addition au temps machine réel de réorganisation d'un temps d'attente (aussi appelé temps gaspillé) correspondant au temps de copie de l'objet juste réorganisé.

A cette fin, l'invention propose une méthode globale de gestion des réorganisations
30 d'un ensemble de bases de données indexées d'un système informatique d'information adaptée à la réorganisation « hors ligne » intégrant la gestion des copies et comportant les phases opérationnelles suivantes (voir figure 2) :

(40) - Création et maintien par mise à jour en continu :

- d'une liste **PRIOREORG** des objets à réorganiser par priorité décroissante
35 fonction de l'état de désorganisation des objets à réorganiser ; et

- d'une liste **PRIOCOPIE** des objets à copier par priorité décroissante fonction de l'ancienneté des dernières copies des objets à copier.

(50) – Surveillance en continu de l'apparition dans le système d'information de tout événement :

- **RAZR=1** pour lequel la sélection du premier ou prochain objet à réorganiser **POR** doit être remise en cause, notamment une fin d'une tâche de réorganisation dans une région de traitement de réorganisation et/ou la libération d'une zone de traitement pour la réorganisation ou la copie ; et/ou
 - **RAZC=1** pour lequel la sélection du premier ou prochain objet à copier **POC** doit être remise en cause, notamment une fin d'une tâche de copie dans une région de traitement de copie, une copie prioritaire en fin de réorganisation et/ou la libération d'une zone de traitement pour la réorganisation ou la copie,
- avec pour tout événement **RAZR=1**, lancement de l'exécution du processus **IDPOR** d'identification du premier objet à réorganiser **POR** ou, si le processus **IDPOR** est en cours d'exécution, interruption et nouveau lancement du processus **IDPOR** ; et/ou pour tout événement **RAZC=1**, lancement de l'exécution du processus **IDPOC** d'identification du premier objet à copier **POC** ou, si le processus **IDPOC** est en cours d'exécution, interruption et nouveau lancement du processus **IDPOC** ;
- le processus **IDPOR** comportant les opérations suivantes

(51) - Etablissement d'un planning rapide de réorganisation **PRR*** à l'aide de la liste **PRIOREORG** et du temps opérationnel de réorganisation résiduel disponible **TRR** dans l'ensemble des régions de réorganisation et d'une liste de sélection **SELECT/OR** d'objets à réorganiser **OR** ordonnés par priorités décroissantes optimisées en fonction du gain de la réorganisation de chaque objet **OR**, ledit gain étant défini comme le produit d'un facteur représentatif du taux de désorganisation d'un objet **OR** par le temps de réorganisation de cet objet;

(52) - Etablissement d'un planning rapide de copie **PRC** à partir de la liste **SELECT/OR**, puis de la liste **PRIOCOPIE** et dans la limite du temps opérationnel de copie résiduel disponible **TRC** dans l'ensemble des régions de copie, d'une liste de sélection **SELECT/OC** en mémoire de type LIFO ou pile d'objets à copier **OC** empilés par priorités décroissantes;

(53) - Etablissement d'un rétroplanning de copie dans lequel pour tout objet courant à réorganiser **OR** extrait de la liste **SELECT/OR** par priorité croissante, on calcule pour chaque région de copie le temps à réserver **TCR** pour la copie de l'objet **OR** après réorganisation ;

(54) - Etablissement d'un rétroplanning de réorganisation dans lequel pour tout objet courant à réorganiser **OR** extrait de la liste **SELECT/OR** par priorité croissante pour favoriser le traitement avancé des objets de plus grandes tailles possibles, on calcule pour chaque région de réorganisation le dernier délai de la région **DDR**,

représentant dans la fenêtre « batch » allouée à la réorganisation « hors ligne » le délai minimal nécessaire à la réorganisation de l'objet courant **OR**, selon une formulation distincte si une copie de l'objet **OR** en fin de réorganisation doit être faite ou non;

- (55) - Identification de la région de traitement **RTR** de l'objet courant à réorganiser
- 5 **OR** par adéquation optimisée entre la durée de traitement disponible dans ladite région **DRR** avec la durée de traitement **DROR** nécessaire pour la réorganisation de l'objet courant **OR** avec $\text{DRR} - \text{DDL}_{\text{minimal}} \leq 0$ sans copie à faire ou avec **valeur absolue de DRR - DDL minimal** et $\text{DROR} + \text{DCOR} \leq \text{DRR}$ si copie à faire;

- (56) - Inscription de l'objet courant à réorganiser **OR** comme prochain objet à
- 10 réorganiser **POR** au planning de la région identifiée en valorisant le pointeur **POR** correspondant à l'adresse de l'objet courant **OR**, en augmentant la durée de temps consommé **DCRR** dans la région du temps de la réorganisation **DROR** et recalculant la durée gaspillée dans la région de réorganisation **DGRR** du fait des copies d'objets **OR** en fin de réorganisation.

- 15 (60) - Lancement, en l'absence de tout événement de remise à zéro **RAZR=1** survenu entre temps, dès qu'une région de traitement de réorganisation **RTR** identifiée est disponible, de la réorganisation de l'objet **POR**; Remise à zéro et nouvelle exécution du processus **IDPOR**; et, le processus **IDPOC** comportant les opérations suivantes

- (51) - Etablissement d'un planning rapide de réorganisation **PRR*** à l'aide de la
- 20 liste **PRIOREORG** et du temps opérationnel de réorganisation résiduel disponible **TRR** et d'une liste de sélection **SELECT/OR** d'objets à réorganiser **OR** ordonnés par priorités décroissantes optimisées en fonction du gain de la réorganisation de chaque objet **OR**, ledit gain étant défini comme le produit d'un facteur représentatif du taux de désorganisation d'un objet **OR** par le temps de réorganisation de cet objet;

- 25 (52) - Etablissement d'un planning rapide de copie **PRC** à partir de la liste **SELECT/OR**, puis de la liste **PRIOCOPIE** et dans la limite du temps opérationnel de copie résiduel disponible **TRC**, d'une liste de sélection **SELECT/OC** en mémoire de type LIFO ou pile d'objets **OC** empilés par priorités décroissantes;

- (53') - Etablissement d'un rétroplanning de copie avec détermination des durées
- 30 de copie disponibles en région copie, dépilage de la liste **SELECT/OC** de haut en bas et recherche d'une région copie par adéquation entre la durée copie disponible et la durée copie de l'objet courant à copier **OC** de façon à copier les objets les plus grands possibles et marquage de l'origine de réorganisation ou non de l'objet à copier; suivi de
- l'identification de la région de traitement **RTC** de l'objet à copier par adéquation entre
- 35 la durée de traitement disponible dans ladite région avec la durée de traitement nécessaire pour la copie de l'objet courant **OC**; suivie de

- l'inscription de l'objet courant **OC** comme le prochain objet à copier **POC** au planning de la région **RTC** identifiée en valorisant le pointeur **POC** correspondant à l'adresse de l'objet **OC** et en diminuant la durée de traitement disponible du temps de la copie de l'objet **OC** retenu comme **POC** ;

- 5 (60')- Lancement, en l'absence de tout évènement de remise à zéro **RAZC=1** survenu entre temps, dès qu'une région de traitement de copie **TRC** identifiée est disponible, de la copie de l'objet **POC**; Remise à zéro et nouvelle exécution du processus **IDPOC**.

Bien entendu sans sortir du cadre de l'invention, il existe dans des variantes de la méthode de l'invention certains cas d'objets à réorganiser **OR** pour lesquels la copie
10 juste après réorganisation n'est pas nécessaire. Dans ces cas et pour ces objets particuliers les phases opérationnelles ci-dessus sont modifiées en conséquence.

En particulier dans une recherche continue d'optimisation, la méthode gestion selon l'invention dans une de ses variantes intègre dans les phases finales de processus **IDPOR** d'un objet **OR** une phase de recherche d'un ou plusieurs objets à réorganiser
15 **OR** sans copie **OR** dans la liste **SELECT/OR** ou à défaut dans la liste **PRIOREORG** susceptibles d'être réorganisés dans la région de traitement **RTR** correspondante pendant l'intervalle du temps d'attente de la copie de l'objet **OR** après réorganisation.

Avantageusement la phase (51) d'Etablissement d'un planning rapide de réorganisation **PRR*** comporte les opérations suivantes (voir figure 3) :

- 20 (511) - Initialisation : dans laquelle sont notamment annulés les indicateurs de sélection des objets **OR**, initialisé le compteur-limiteur de remises en causes, les temps résiduels de copie **TRC** et de réorganisation **TRR** et la durée minimale **E** de réorganisation et copie d'un objet **OR** ;

(512) - Boucle sur objets : dans laquelle notamment pour tout objet réel **OR** de la liste
25 **PRIOREORG** pris dans l'ordre décroissant jusqu'à ce que **TRR < E**, une marque de 'sélection possible' de l'objet est posée;

(513) - Elimination des intersections d'objets : dans laquelle notamment sont éliminés de la 'sélection possible' les objets réels **OR** parties d'espaces de tables également appelés objets globaux déjà sélectionnés en tout ou partie;

- 30 (514) - Remise en cause des choix précédents : dans laquelle notamment sont éliminés de la 'sélection possible' les objets **OR** dont le temps total de réorganisation et copie est supérieur au temps de réorganisation résiduel **TTR** et à défaut dans laquelle l'objet courant **OR** et les objets **OR** déjà sélectionnés dans la liste **SELECT/OR** sont soumis à un processus d'optimisation en fonction du gain de
35 réorganisation avec remplacement d'un objet déjà sélectionné par l'objet en cours avec essais, dans la limite du terme du compteur-limiteur, de combinaisons successives tendant à retenir pour être sélectionné au final dans **SELECT/OR** des

objets de plus grand gain de réorganisation possible et à les associer aux autres objets **OR** de la 'sélection possible' de façon à ce que le temps total des réorganisations de tous les objets **OR** retenus soit inférieur mais le plus proche possible du temps résiduel de réorganisation **TRR**.

5 (515) - Vérification de la suffisance du temps copie disponible dans le temps résiduel de copie **TRC** pour chaque objet **OR** en passe d'être finalement sélectionné pour la réorganisation compte tenu de la priorité de réorganisation de l'objet **OR** par rapport aux objets à copier en préalable du fait de leurs priorités copie supérieures.

10 (516) - Sélection de l'objet courant dans la liste **SELECT/OR** : dans laquelle notamment sont positionnés les indicateurs de l'objet courant et de l'objet global associé, annulés éventuellement les indicateurs du ou des objets **OR** désélectionnés et diminuer **TRR** et **TRC** des temps respectifs de réorganisation et de copie.

15 Encore avantageusement, ladite phase (52) d'établissement d'un planning rapide de copie **PRC** comporte les opérations suivantes (voir figure 4) :

(521) - Initialisation : au sortir d'un planning de réorganisation **PRR*** et dans laquelle sont notamment vidées les piles de sélection de copie **PSC1** et **PSC2**, initialisés le compteur-limiteur de remises en causes, le temps résiduel de copie **TRC** et la durée
20 minimale **EC** de copie d'un objet à copier **OC** ;

(522) – Inspection des régions de réorganisation : Pour toutes les régions, si région active, et si image copie de l'objet en cours de réorganisation à faire et/ou si image copie de tout objet **OR** de la liste **SELECT/OR** pris dans l'ordre des priorités décroissantes à faire: pour tout sous-objet susceptible d'être copié séparément,
25 notamment une partition, soustraction du temps de copie du sous-objet de **TRC** et placement du sous-objet en haut de la première pile **PSC1** ;

(523) – traitement de la liste des objets à copier **PRIOCOPIE** : dans laquelle notamment pour tout objet à copier **OC** de la liste **PRIOCOPIE** pris dans l'ordre décroissant jusqu'à ce que **TRC < EC** et absent de **PSC1**, soustraction du temps de
30 copie du sous-objet **OC** de **TRC** et placement du sous-objet **OC** en haut de la seconde pile **PSC2** ,

(524) – Empilage des piles de sélection : dans laquelle la pile **PSC1** est superposée à la pile **PSC2** pour constituer la liste **SELECT/OC**.

Encore avantageusement, ladite phase (53) d'établissement d'un
35 rétroplanning de copie dans le processus **IDPOR** comporte les opérations suivantes :

(531) Initialisation des régions de copie : dans laquelle au sortir d'un planning de réorganisation pour toutes les régions de copie, la durée consommée de la région **DCRC** est mise à zéro,

(532) Détermination du temps de copie à réserver **TCR** : dans laquelle pour tout objet réel de la liste **SELECT/OR** parcourue dans l'ordre des priorités croissantes et pris comme objet à réorganiser courant **OR**,

-le temps à réserver pour copier **TCR** l'objet réel courant **OR** est mis à zéro ;

-Si l'image copie de tout objet **OR** de la liste **SELECT/OR** pris dans l'ordre des priorités croissantes doit être faite: pour tout sous-objet susceptible d'être copié séparément, notamment partition, on recherche une région copie dont la durée consommée **DCRC** est minimale, on ajoute le temps de copie du sous-objet à la durée consommée de la région courante **DCRC** et on valorise le temps à réserver **TCR** pour copie de l'objet courant = **maximum** (du temps à réserver **TCR** pour copie de l'objet courant et de la durée consommée dans la région **DCRC**).

Encore avantageusement, la phase de rétroplanning de réorganisation (54) dans le processus **IDPOR** comporte les opérations suivantes :

(541) – Initialisation : dans laquelle pour toute région de réorganisation,

- Durée de la région **DRR** = Durée de la fenêtre 'batch' **DFB** moins l'heure relative estimée de fin du traitement de réorganisation en cours dans la région **HRFR**

- Durée consommée dans la région **DCRR** = 0

- Durée gaspillée dans la région **DGRR** = 0

- Pointeur premier objet à réorganiser de la région courante = 0

(542) – Détermination du dernier délai de la région de réorganisation **DDRR** : dans laquelle pour tout objet réel **OR** sélectionné de la liste **SELECT/OR** dans l'ordre des priorités croissantes :

Pour chaque région si l'image copie de l'objet courant **OR** à faire, calcul du dernier délai de la région **DDRR** = durée de réorganisation de l'objet courant **DROR** + **max** [temps à réserver pour copie de l'objet courant **TCR**, (durée consommée de la région **DCRR** + durée gaspillée de la région **DGRR**)] ;

- Sinon, calcul du dernier délai de la région **DDR** = durée consommée de la région **DCRR** + **max** [(durée de réorganisation de l'objet courant **DROR**, durée gaspillée de la région **DGRR**)].

Encore avantageusement la phase d'identification de la région de réorganisation **RTR** (55) comporte les opérations suivantes : Selon le type d'objet courant à réorganiser **OR**

(551) – Recherche de la région de traitement de réorganisation susceptible d'accueillir de façon optimale un objet à réorganiser **OR** sans copie : comportant

- la recherche d'une région dont le dernier délai **DDRR** est **minimal**, et si dernier délai de la région courante **DDR** est **supérieur** à la durée de la région courante **DRR**, la recherche d'une région pour laquelle : [durée de la région courante **DRR** moins le dernier délai de la région courante **DDRR**] est **minimale et positive ou nulle** ; sinon
- 5 (552) – Recherche de la région de réorganisation susceptible d'accueillir de façon optimale un objet à réorganiser **OR** avec copie : comportant
 - la recherche d'une région telle que : [durée de la région **DRR** moins dernier délai de la région courante **DDRR**] est **minimale en valeur absolue**, et [durée de réorganisation **DROR** + temps de copie objet courant **TCR**] \leq durée de la région **DRR**.
- 10 Encore avantageusement la phase d'inscription de l'objet **OR** au planning de la région de réorganisation **RTR** (56) comporte les opérations suivantes :
 - (561) – Inscription de l'objet courant **OR** comme prochain objet à réorganiser **POR**: dans laquelle si image copie de l'objet courant **OR** à faire ou durée gaspillée de la région courante = 0, on valorise le **pointeur premier objet à réorganiser POR de la**
 - 15 **région courante = adresse de l'objet courant**,
 - (562) - mise à jour **DCRR** et **DGRR** : dans laquelle on ajoute la durée de réorganisation de l'objet courant **DROC** à la durée consommée de la région courante **DCRR**, et on valorise la durée gaspillée de la région courante **DGRR** = dernier délai de la région courante **DDRR** – durée consommée de la région courante **DCRR**.
- 20 Toujours avantageusement la phase (53') d'établissement d'un rétroplanning de copie dans le processus **IDPOC** comporte les opérations suivantes:
 - (531') Initialisation des régions de copie : dans laquelle pour toutes les régions,
 - la durée de région **DRCOP** est valorisée à la durée fenêtre « batch » **DFB** moins l'heure relative estimée de fin de traitement copie en cours **HRFC**,
 - 25 -la durée consommée de la région **DCRCOP** est mise à zéro,
 - le pointeur du premier objet à copier **POC** de la région courante est mis à zéro
 - (532') Dépilage et identification du prochain objet à copier **POC** : Dans la laquelle,
 - Pour tout objet de la pile **PSC1/PSC2** parcourue de haut en bas un objet courant à copier **OC** est extrait lorsque le temps de copie de l'objet courant **TCOC** \leq **DRCOP** la
 - 30 plus longue ;
 - On cherche pour identifier la région de traitement copie une région dont la durée consommée **DCRC** est minimale ; dans l'affirmative si le temps de copie de l'objet courant **TCOC** $>$ **DRCOP** – **DCRCOP** on cherche une région pour laquelle **DRCOP** – **DCRCOP** – **TCOC** minimale, positive ou nulle,
 - 35 Si aucune région ne convient, on choisit une région à copier pour laquelle **DRCOP** – **DCRCOP** – **TCOC** est négative et minimale en valeur absolue et pour laquelle **TCOC** \leq **DRCOP** ;

-Si l'objet courant appartient à la seconde pile **PSC2**, soit un objet à copier non issu d'une réorganisation, le pointeur de l'objet **POC** pour la région courante choisie est valorisé à l'adresse de l'objet courant **OC** de façon à inscrire l'objet **POC** au planning de la région de traitement copie **RTC** identifiée ; et

- 5 -Pour terminer, on ajoute la durée copie **TCOC** à la durée consommée de la région courante **DCRCOP**.

Selon les diverses variantes de mise en œuvre de la méthodes de gestion de réorganisations « hors ligne » dans des régions de traitement il est possible de choisir les évènements internes au système remettant en cause la sélection des objets **POR** ou **POC** les plus appropriés aux processus de sélection utilisés.

10 En particulier les événements internes au système **RAZR=1** et **RARC=1** entraînant la remise en cause de la sélection des objets **POR** ou **POC** sont notamment constitués respectivement par la fin d'une tâche de réorganisation dans une région de traitement de réorganisation pour les objets **POR** ou par la fin d'une tâche de copie dans une région de traitement de copie pour les objets **POC**, une copie prioritaire en fin de réorganisation et/ou la libération d'une zone de traitement pour les objets **POR** et **POC**.

Parallèlement ou non à la mise en œuvre la méthode adaptée aux réorganisations « hors ligne », l'invention propose une méthode de gestion des réorganisations dans un ensemble de bases de données indexées d'un système d'information adaptée à la réorganisation « en ligne » et caractérisée en ce qu'elle comporte au moins les phases opérationnelles suivantes:

- 20 - une phase d'analyse instantanée d'au moins un objet à réorganiser **OR** parmi les objets susceptibles de réorganisation, notamment bases de données, partitions et/ou index, et d'estimation du surcoût associé au niveau de désorganisation de l'objet à réorganiser **OR**;
- 25 - une phase d'estimation instantanée du coût de la réorganisation en ligne en fonction de la taille dudit objet à réorganiser **OR** et du taux de désorganisation dudit objet **OR**; et
- 30 - une phase de détermination du seuil **Ds** de désorganisation minimal de l'objet à réorganiser **OR** au dessus duquel seuil la réorganisation « en ligne » peut être lancée pour cet objet **OR** et le, cas échéant, du lancement effectif de cette réorganisation en ligne, le seuil **Ds** correspondant sensiblement au minimum du coût total du surcoût estimé de la désorganisation de l'objet **OR** et du coût estimé de la réorganisation pour l'objet **OR** considéré.

35 Selon une première variante, la méthode de gestion des réorganisations « on line » est caractérisée en ce que le lancement de la réorganisation « en ligne » est retardé pour attendre une fenêtre temporelle d'activité réduite de la base de données

concernée. Selon une variante particulière de l'invention la méthode de gestion incorpore une phase de calcul de l'heure optimale de lancement d'une réorganisation «en ligne» (on line) en fonction de l'activité réelle ou prévue du système d'information correspondant.

- 5 D'une façon plus globale les méthodes des réorganisations « en ligne » et « hors ligne » sont combinées de façon à ce que pour un objet **OR** sélectionné comme prochain objet à réorganiser **POR** la priorité est donnée pour la réorganisation « hors ligne », l'intervention de la réorganisation « en ligne » n'étant requise qu'après dépassement du seuil **Ds**. Il est ainsi possible de soulager la réorganisation « hors
- 10 ligne » par des réorganisations « en ligne » lorsque ces dernières ne sont pas trop pénalisantes au niveau de la qualité des services en ligne.

- L'invention concerne également une méthode de gestion des réorganisations « en ligne » et « hors ligne » dans un ensemble de base de données indexées caractérisée en ce que pour un même objet **POR** la priorité est donnée pour la réorganisation « hors
- 15 ligne », l'intervention de la réorganisation « en ligne » n'étant requise qu'après dépassement du seuil **Ds** pour l'objet **POR** considéré.

- Selon une première version de la méthode de gestion des réorganisations « on line » utilisée pour la réorganisation des bases de données indexées de type espace de tables, la méthode est caractérisée en ce que pour un objet à réorganiser **OR**, le seuil
- 20 **Ds** est défini par une approximation donnée par la formule **F5'** suivante:

$$R = Ds^2 r Co^* / 2.l.c = 1$$

lorsque que **U**, moyenne horaire du nombre de mises à jour sur l'objet **OR**, est faible, et dans laquelle :

r désigne le nombre moyen de lignes ou clé-RID par page de l'objet entier,

- 25 **l** désigne le nombre d'insertions de lignes (ou de clé-RID pour les index) à l'heure dans l'objet **OR**,

Co* désigne le surcoût horaire lorsque l'objet est totalement désorganisé

c désigne un paramètre machine, le rapport du temps d'accès **E/S** par la taille de l'objet **Nbp**.

- 30 Selon une variante de la méthode de gestion des réorganisations « on line » précédente et utilisée pour la réorganisation des index, le seuil **Dsl** est défini pour un index par une approximation donnée par la formule **F7'** suivante :

$$R = Dsl^2 r (t^*+1)Nbp.V / [2(V+3)] = 1$$

- lorsque que **U**, moyenne horaire du nombre de mises à jour sur l'index, est faible, et
- 35 dans laquelle :

Ts désigne le temps de déplacement de bras, **Tl** le temps de latence (1/2 tour du bras).

$c = (T_s + 3 \cdot T_l) / N_{bp}$ où T_s en E/S désigne le temps de déplacement du bras du disque, et T_l le temps de latence (1/2 tour du bras) et $V = T_s / T_l$.

t^* désigne le taux de relecture de l'objet, c'est à dire le nombre moyen de fois qu'une entrée sera relue.

5 Selon une autre variante de la méthode de gestion des réorganisations « en ligne » précédente et utilisée pour la réorganisation des bases de données indexées de type espace de tables monotable à n index, caractérisé en ce que le seuil DsT est défini pour un espace monotable par une approximation donnée par la formule **F13'** suivante :

$$10 \quad R = r \cdot N_{bp} [DsT^2 \cdot L(V+1) + V \cdot S(DsT^2 - b_i^2) (t^*i+1) / (2V+6)] = 1$$

lorsque que U , moyenne horaire du nombre de mises à jour sur l'espace monotable, est faible, et dans laquelle :

L désigne la proportion de lignes accédées par parcours séquentiel sur le nombre de lignes créées,

15 S représente le signe mathématique Sigma (somme des expressions $f(x_i)$ pour i index allant de 1 à n),

b_i représentent les valeurs instantanées de D (taux de désorganisation de l'espace monotable) lors des dernières réorganisations des n index, et

t^*i représentent les taux d'accès directs au travers de chaque index i .

20 L'invention concerne également l'application de la méthode de gestion des réorganisations dans un ensemble de bases de données indexées d'un système informatique d'information présentée ci-dessus dans toutes ses variantes à la réorganisation des bases de données indexées de type espace de tables défini dans la description ci-après.

25 L'invention concerne également les systèmes informatiques de type systèmes d'information comportant un ensemble de bases de données indexées et les moyens adaptés pour la mise en œuvre de la méthode de gestion des réorganisations présentée ci-avant, de base de données indexées, notamment de bases de données de type espace de tables et/ou d'index.

30 D'autres caractéristiques et avantages de la méthode de gestion globale des réorganisations de bases de données indexées et de ses applications à la réorganisation des bases de données indexées notamment du type espace de tables selon la présente invention apparaîtront à la lecture de la description qui va suivre en référence aux dessins ci-joints d'un mode de mise en œuvre préférentiel de l'invention
35 et présenté à titre d'exemple non limitatif.

PRESENTATION SOMMAIRE DES DESSINS

- la figure 1 représente un organigramme correspondant à la méthode de gestion des réorganisations d'un ensemble de bases de données indexées, notamment de type espace de tables, selon l'invention;
- la figure 2 représente un organigramme correspondant à la méthode de gestion des réorganisations et des copies d'un ensemble de bases de données indexées, notamment de type espace de tables, selon l'invention;
- la figure 3 représente un organigramme détaillant les opérations de la phase d'établissement d'un planning rapide de réorganisation de la méthode de la figure 2 ;
- la figure 4 représente un organigramme détaillant les opérations de la phase d'établissement d'un planning rapide de copie de la méthode de la figure 2 ;
- la figure 5A représente un organigramme correspondant à une méthode de lecture d'une base de données indexée, notamment en vue de la réorganisation de la base de données, utilisable dans le cadre de l'invention;
- la figure 5B représente un organigramme correspondant à une méthode de lecture de l'index de la base de données indexée, notamment en vue de la réorganisation de l'index, utilisable dans le cadre de l'invention;
- la figure 6A représente un organigramme correspondant à une variante partielle de la méthode de la figure 5A appliquée à la lecture et au déchargement de l'index d'une base de données indexée de type espace de tables, notamment en vue de la réorganisation de la base de données indexée, utilisable dans le cadre de l'invention;
- la figure 6B représente un organigramme correspondant à une variante partielle de la méthode de la figure 5B appliquée à la lecture et au déchargement de l'index d'une base de données indexée de type espace de tables, notamment en vue de la réorganisation de l'index de la base de données indexée, utilisable dans le cadre de l'invention;
- la figure 7 représente un organigramme correspondant à une variante partielle de la méthode appliquée à la lecture et au déchargement de la base de données indexée de type espace de tables objet de la figure 6A, utilisable dans le cadre de l'invention;
- la figure 8A représente un organigramme correspondant une méthode de réorganisation d'une base de données indexée de type espace de tables utilisant les méthodes de lecture et de déchargement objet des figures 6A et 7, utilisable dans le cadre de l'invention; et
- la figure 8B représente un organigramme correspondant une méthode de réorganisation de l'index de la base de données indexée de type espace de tables

objet de la figure 8A, méthode utilisant la méthode de lecture et de déchargement objet de la figure 6B, utilisable dans le cadre de l'invention.

La description de la méthode de gestion globale de réorganisation d'un ensemble de bases de données indexées qui va suivre, présentée à titre d'exemple non limitatif, se rapporte à une application particulière de l'invention aux bases de données indexées structurées en espaces de tables ('tablespace' en anglais), susceptibles d'atteindre de très grandes tailles (jusqu'à quelques milliards de lignes) et pour lesquelles le maintien d'une bonne organisation, notamment par une réorganisation rapide et efficace, est capital du point de vue performances.

10 Exemple de bases de données indexées, les bases de données de type espace de table :

Une base de donnée indexée structurée en espace de tables est composée d'un ou plusieurs fichiers formés de blocs de taille unique ou pages numérotées en séquence à partir de zéro. Du point de vue logique on considère qu'il n'existe qu'un seul fichier composé de la succession de chacun des fichiers physiques numérotés **1, 2**, etc... .

15 Parmi ces pages on trouve les pages de contrôle et les pages de données adaptées pour contenir des lignes, chaque ligne étant identifiée de façon unique par son RID composé du numéro de page et du numéro d'ordre à l'intérieur de la page. Chaque ligne est composée de champs de données appelées colonnes. Une clé est une liste de colonnes ayant un ordre croissant ou décroissant. A chaque clé est associée un index.

20 Un espace de tables peut contenir une ou plusieurs tables présentant chacune des colonnes spécifiques. Toutefois les espaces de tables multi-tables peuvent être traités comme les espaces de tables mono-table à condition de traiter les tables et leurs index en séquence.

25 Les espace de tables peuvent être partitionnés (divisés en plusieurs parties ou partitions). Chaque partition étant alors identique au fichier physique portant son numéro.

30 A la création de la base de données, il est prévu un espace libre de données et un espace libre d'index constitués chacun de pages vides réparties dans l'espace de tables et dans l'index; le but de cette répartition de pages vides entre certaines pages de données et index est de permettre des ajouts et autres mises à jour de lignes de données ou de clés d'index à proximité des pages correspondantes et de limiter ainsi les désorganisations trop importantes.

On considère qu'un espace de tables est bien organisé lorsque:

- 35 - l'espace libre de données (pages vides) est distribué dans l'espace de tables selon les paramètres d'origine ;
- les lignes sont rangées par ordre de RID dans le même ordre que la première clé dite clé 'primaire'.

L'index est composé comme l'espace de tables d'un ou plusieurs fichiers formés de pages. Les pages sont de plusieurs type, en particulier pages de contrôle, pages 'feuilles' et pages de niveau supérieur. Les pages feuilles contiennent des clés et pour chaque clé le ou les RID de la ou des lignes d'espace de tables dont la clé correspondante a la même valeur. Ainsi à chaque ligne d'espace de tables est associé un couple clé-RID de page feuille de l'index.

Chaque page feuille contient un ensemble contigu de couples clé-RID, c'est à dire qu'il n'existe aucun couple extérieur à la page qui soit compris entre deux couples de la page. Les pages feuilles sont ainsi logiquement ordonnées par l'un quelconque de leurs éléments.

En cas de partitions d'espace de tables, il en sera de même pour son index primaire, une partition d'index étant associée à chaque partition de l'espace de tables.

On considère qu'un index est bien organisé lorsque:

- l'espace libre est distribué dans l'espace de tables selon les paramètres d'origine ;
- les pages feuilles sont rangées par numéros dans le même ordre que les clés.

La définition de base de données indexée de type 'espace de tables' donnée ci-dessus restera valable pour la suite de l'exposé.

D'une façon générale la réorganisation unitaire d'une base de données (et de son index) consiste à décharger la base de données, à ranger ses lignes d'information dans un ordre souhaité, par exemple l'ordre alphabétique, puis à recharger les lignes une fois rangées dans la base de données en mettant à jour les index, si possible de façon réorganisée.

Exemple de réorganisation « hors-ligne » de bases de données indexées, notamment de type espace de tables :

Toujours à titre d'exemple non limitatif, on utilise pour la réorganisation unitaire des bases de données indexées structurée en lignes et notamment les bases de type espace de tables un mode de lecture (déchargement) et un mode de réorganisation permettant de réduire sensiblement le temps nécessaire à la lecture ordonnée de la base de données indexée (lecture et éventuelle mise en ordre continue ou discontinue) et/ou à la réorganisation de la base de données indexée.

En particulier on utilise de façon générale comme méthode de lecture de bases de données indexée dans laquelle les données apparaissent sous forme de lignes de données, la méthode **LEC/BD/REORG/BD**: comportant les opérations suivantes (voir figure 5A):

[101] - Lecture simultanée et séquentielle de l'index et de lignes de la base de données, BDD, avec le chargement des index dans une première mémoire tampon MX et le chargement des lignes dans une seconde mémoire tampon MT,

[102] - Lecture des index de la mémoire tampon MX vers mémoire tampon MT de façon ordonnée,

[103] - Recherche dans la mémoire MT de la présence d'une ligne dont l'identifiant est donné par la lecture ordonnée de l'index,

5 **[103A]** - Dans l'affirmative, extraction de la ligne vers un fichier de lignes ordonnées logiquement FT1,

[103B] - Dans la négative, vidange partielle de la mémoire tampon vers un fichier FT2 de lignes à trier avec élimination de la mémoire tampon MT d'au moins la page dont le numéro est le plus faible,

10 **[104]** - Poursuite de la lecture jusqu'à la fin de la base de données, les lignes restantes n'ayant pu être extraites de façon ordonnée étant dirigées vers le fichier FT2.

Appliquée à la réorganisation de la base de données, la méthode de lecture **LEC/BD/REORG/BD** permet, après le tri du fichier FT2 pour donner FT'2, de combiner les deux fichiers ordonnés FT1 et FT'2 en un fichier virtuel FTV servant d'entrée à la phase de rechargement de la réorganisation de la base de données.

15 Il est à noter que par fichier virtuel on entend un fichier qui n'a pas d'existence réelle, car ce fichier est le produit théorique de la fusion virtuelle de deux de fichiers, mais qui peut être utilisé au rechargement comme un fichier réel dans la mesure où le programme de rechargement lit à la fois sur le fichier FT1 et à la sortie FT'2 du tri
20 (elle-même un fichier virtuel généralement nommé « exit de tri ») et opère la fusion ligne à ligne au lieu de lire un seul fichier. L'expression fichier virtuel désigne plus généralement une entité sur laquelle on peut opérer les mêmes fonctions que sur un fichier séquentiel classique (ouverture, lecture ou écriture d'un enregistrement, fermeture) et qui du point de vue de ces fonctions se comporte exactement comme un
25 fichier séquentiel classique.

Ainsi grâce à la génération, en continu et par simple extraction, d'un fichier intermédiaire de lignes ordonnées logiquement FT1 pendant l'opération de lecture séquentielle, le nombre de lignes à trier présentes en en fin de parcours dans le fichier FT2 est réduit de façon substantielle d'où un gain de temps sensible tant pour
30 l'opération de lecture de la base de données que pour le traitement complet de la réorganisation de la base de données.

La méthode de lecture de bases de données **LEC/BD/REORG/BD** permet à partir d'une lecture de type séquentielle de la base de données d'obtenir un fichier virtuel ordonné des lignes de données tout en minimisant l'étape de tri effectuée sur un fichier
35 intermédiaire partiel.

De façon analogue on utilise comme méthode de lecture d'index de bases de données indexée dans laquelle les données apparaissent sous forme de lignes de

données, la méthode **LEC/IND/REORG/IND** comportant les opérations suivantes (voir figure 5B):

- [101X] - Lecture séquentielle de l'index avec le chargement des index dans une mémoire tampon **MX**,
- 5 [102X] - Vérification de la position du couple clé/ligne de plus bas numéro d'ordre (clé/ligne **INF**) présent dans mémoire tampon **MX** par rapport à la position du dernier couple traité, en particulier si le numéro d'ordre du couple clé/ligne **INF** est supérieur au numéro d'ordre du dernier couple traité
- [103XA] - Dans l'affirmative, extraction du couple clé/ligne **INF** vers un fichier de
- 10 couples d'index logiquement ordonnés **FX1**,
- [103XB] - Dans la négative, vidange partielle de la mémoire tampon **MX** vers un fichier **FX2** de couples à trier avec élimination de la mémoire tampon **MT** d'au moins le couple clé/ligne dont le numéro est le plus faible,
- [104X] - Poursuite de la lecture jusqu'à la fin de la base de données, les couples
- 15 clé/ligne restant n'ayant pu être extraits de façon ordonnée étant dirigés vers le fichier **FX2**.

Ainsi l'opération de lecture ordonnée de l'index, continue ou discontinue en l'absence de certaines valeurs de clé d'index, permet l'extraction directe d'un fichier d'index logiquement ordonné **FX1** et d'un fichier d'index à trier **FX2**. Après le tri du

20 fichier **FX2** pour donner **FX'2**, les deux fichiers ordonnés **FX1** et **FX'2** sont fusionnés en un fichier virtuel **FXV** servant d'entrée à la phase de rechargement de la réorganisation de l'index.

Dans la pratique on arrive à constituer des fichiers **FT1** contenant plus de 90% des lignes lues. D'autre part la durée d'un tri est plus que proportionnelle à la quantité

25 d'information en entrée, elle se trouve divisée par au moins 10. En ce qui concerne le temps de lecture, celui-ci reste bref car l'opération de lecture est réalisée de façon séquentielle et simultanée sur la base de données et sur l'index.

Selon un premier mode de mise en oeuvre de la méthode de lecture **LEC/BD/REORG/BD** et applicable à une base de données indexée de type 'espace de

30 tables', le processus de lecture et de déchargement de l'index à l'aide la mémoire tampon **MX** est réalisé par la méthode **LEC/IND/REORG/TABLE** comportant les opérations suivantes (voir: figure 6A) :

- [201] - Autant de fois que nécessaire jusqu'à avoir atteint la fin de l'index et vidé la mémoire tampon **MX**, lecture séquentielle des pages de l'index en ne conservant que
- 35 les pages d'index contenant au moins une clé, jusqu'à remplir la mémoire tampon ou atteindre la fin de l'index ; cette lecture se faisant dans les emplacements non utilisés de la mémoire tampon ou contenant une page marquée invalide;

[202] – Recherche et remplacement des pages invalides dans la mémoire tampon de l'index à partir de la première page de la mémoire tampon **MX**,

[206A] - Vérification dans chaque page courante de la mémoire tampon :

5 - de l'absence d'au moins un couple clé-RID supérieur au dernier couple clé-RID traité en séquence logique (écrit dans le fichier index trié **FX1** ou envoyé en lecture d'espace de tables/base de données); ou

10 - de l'exactitude de l'expression $PXC \leq PXL - DMA$ où **PXC** représente le numéro dans l'index de la page courante de la mémoire tampon **MX**, **PXL** le numéro dans l'index de la dernière page d'index lue et **DMA** la différence maximale autorisée entre pages d'index en cours de lecture;

[206B] - Dans l'affirmative d'au moins une des deux propositions, invalidation de la page courante **PXC** et retour à l'opération [201] pour lecture d'une nouvelle page index en séquence;

15 [203] – Recherche de la page valide **PXmin** de la mémoire tampon **MX** qui contient la clé la plus basse **CLE INF** de toute la mémoire ou en cas d'égalité de clé, le RID correspondant le plus bas;

[204] – Traitement de la page **PXmin** en séquence logique comportant les opérations suivantes :

20 [208A] – Le cas échéant, lecture par ordre croissant de chacun des couples clé-RID de cette page **PXmin**;

[208B] - Envoi de chacun des RID de cette page **PXmin** dans l'ordre croissant des couples clé-RID vers le traitement de lecture de la base de données de type espace de tables;

25 [204C] - Invalidation de la page **PXmin** traitée et retour à l'opération [201] pour lecture d'une nouvelle page index en séquence.

On notera que l'opération **208B** constitue un déchargement simplifié d'index, dans lequel seuls les RID des pages extraites en séquences logiques sont transmis et utilisés pour la réorganisation de l'espace de tables.

30 En particulier la méthode de lecture et de déchargement **LEC/TABLE/REORG/TABLE** de l'espace de tables à l'aide la mémoire tampon **MT** est réalisée par les opérations suivantes (voir figure 7):

35 [301] – Autant de fois que nécessaire jusqu'à avoir atteint la fin de l'espace de tables et vidé la mémoire tampon **MT**, lecture séquentielle des pages de l'espace de tables ou de la partition à partir du début en ne conservant que les pages de données contenant au moins une ligne, jusqu'à remplir la mémoire tampon **MT** ou atteindre la fin de l'espace de tables;

[302] - Pour chaque RID présent provenant de la lecture de l'index (voir opération 208B):

- si le RID désigne un numéro de page **PTRID** inférieur au minimum **PTmin** des numéros de pages dans la mémoire tampon **MT**, le RID est ignoré parce que la page déjà été traitée; retour au début de l'opération **[302]**;

- si le RID désigne un numéro de page **PTRID** supérieur au maximum des numéros de pages **PTmax** dans la mémoire **MT**, la page n'a pas encore été lue et on effectue **[303]**; si non on effectue **[304]** ;

- si le RID est absent on effectue **[305]** ;

[303] – Si $\text{PTRID} - \text{PTmax} \leq \text{INC}$, où l'incrément autorisé **INC** représente le nombre maximal de pages de l'espace de tables qui peuvent être lues pour retrouver un RID donné, on effectue **[303A]** autant de fois que nécessaire jusqu'à obtenir la page portant le numéro de RID ; à défaut on effectue **[303B]**;

[303A]- Vidange du contenu de la page de la mémoire **MT** ayant le plus petit numéro de page **PTmin** en envoyant les lignes correspondantes dans le fichier tri-espace de tables **FT2**; lecture à l'emplacement de **Pmin**, de la page suivante à la page **PTmax** de la mémoire tampon **MT** et on recommence jusqu'à obtenir une page de données contenant au moins une ligne; lorsque la dernière page introduite dans la mémoire **MT** porte bien le numéro de RID, on effectue l'opération **[304]**;

[303B]- Si $\text{PTRID} - \text{PTmax} > \text{INC}$, et si $\text{PTRID}(n) - \text{PTRID}(n-1) < \text{PRM}$, où le 'paramètre de proximité' **PRM** représente l'écart maximal des numéros de pages de données que deux RID consécutifs provenant de la lecture d'index ne devraient pas dépasser, sauf en cas de trou dans la séquence, les deux pages **PTRID(n)** et **PTRID(n-1)** sont probablement voisines et on double la valeur de l'incrément autorisé **INC**, à défaut **INC** reste inchangé; dans les deux cas on ignore le RID avec un retour à l'opération **[302]** ;

[304]- Si le RID désigne une page **PTRID** présente dans la mémoire tampon **MT** (y compris après la lecture **[303A]**), il suffit pour cela que le numéro de page soit entre les bornes puisque que la mémoire tampon **MT** contient des pages non vides contiguës, on écrit la ligne correspondante sur le fichier des lignes ordonnées **FT1** et on divise l'incrément autorisé **INC** par 2, pour autant que le résultat ne soit pas en dessous de sa valeur minimale ; retour en **[302]**

[305] – une fois traité tous les RID disponibles, on vide vers le fichier **FT2** toutes les pages restantes de la mémoire tampon **MT** et éventuellement toutes les pages de l'espace de tables encore non lues.

La méthode de lecture et de déchargement ci-dessus **LEC/TABLE/REORG/TABLE** est appliquée de la façon suivante pour la réorganisation des bases de données indexées de type espace de tables.

En particulier on utilise comme méthode de réorganisation unitaire d'une base de données indexée organisée en espace de tables, la méthode **REORG/TABLE** comportant les opérations suivantes (voir figure 8A):

- [401]** – Lecture séquentielle et simultanée du contenu de l'index et du contenu de l'espace de tables et création des fichiers de lignes de données **FT1** et **FT2**;
- [402]** – Si nécessaire, tri du fichier **FT2** par ordre de clé primaire croissante et éventuellement par ordre de table croissante s'il y a plusieurs tables;
- [403]** – Fusion virtuelle des fichiers de lignes **FT1** et **FT2**;
- [404]** – Rechargement du ou des fichiers contenant l'espace de tables avec les lignes ainsi ordonnées ou triées en respectant les paramètres d'espace libre ;
- [405]** – Rechargement de l'index primaire de chaque table dans le même ordre en prenant en compte le nouvel emplacement (RID) de chaque ligne;
- [406]** – Mise à jour des autres index soit en modifiant les RID soit en les rechargeant après tri des informations de clé et de RID.

En ce qui concerne la réorganisation de l'index ou d'une partition d'index de la base de données indexée, celle-ci est susceptible, entre autres, de mettre en œuvre (en variante de la méthode de lecture et de déchargement d'index illustrée à la figure 6A) la méthode de lecture et déchargement d'index de base de données indexée de type espace de tables à l'aide la mémoire tampon **MX** et réalisée par la méthode **LEC/IND/REORG/IND(ET)** comportant les opérations suivantes (voir figure 6B) :

- [201]** – Autant de fois que nécessaire jusqu'à avoir atteint la fin de l'index et vidé la mémoire tampon **MX**, lecture séquentielle des pages de l'index en ne conservant que les pages d'index contenant au moins une clé, jusqu'à remplir la mémoire tampon ou atteindre la fin de l'index ; cette lecture se faisant dans les emplacements non utilisés de la mémoire tampon ou contenant une page marquée invalide;
- [202]** – Recherche et remplacement des pages invalides dans la mémoire tampon de l'index à partir de la première page de la mémoire tampon **MX**,
- [206A]** – Vérification dans chaque page courante de la mémoire tampon :
 - de l'absence d'au moins un couple clé-RID supérieur au dernier couple clé-RID traité en séquence logique (écrit dans le fichier index trié **FX1** ou envoyé en lecture d'espace de tables/base de données); ou
 - de l'exactitude de l'expression $PXC \leq PXL - DMA$ où **PXC** représente le numéro dans l'index de la page courante de la mémoire tampon **MX**, **PXL** le numéro dans l'index de

la dernière page d'index lue et **DMA** la différence maximale autorisée entre pages d'index en cours de lecture;

(202'B) - Dans l'affirmative d'au moins une des deux propositions, écriture des couples clés - RID dans le fichier tri-index **FX2**, invalidation de la page courante **PXC** et retour à l'opération (201) pour lecture d'une nouvelle page index en séquence; à défaut on exécute (203) ;

[203] - Recherche de la page valide **PXmin** de la mémoire tampon **MX** qui contient la clé la plus basse de toute la mémoire **MX** ou en cas d'égalité de clé, le RID correspondant le plus bas;

[204] - Traitement de la page **PXmin** en séquence logique comportant les opérations suivantes :

[204'A] - Lecture par ordre croissant de chacun des couples clé-RID de cette page **PXmin** et écriture directe de ces couples clé-RID dans le fichier index ordonné **FX1**;

[204C] - Invalidation de la page **PXmin** traitée et retour à l'opération (201) pour lecture d'une nouvelle page index en séquence.

Ainsi le déchargement de l'index selon cette dernière méthode conduit à la création d'un fichier virtuel **FXV** ordonné des couples clé-RID et la méthode **REORG/IND(ET)** de réorganisation de l'index de la base de données indexée et organisée en espace de tables comporte alors les opérations suivantes (voir figure 8B):

[401X] - Lecture du contenu de l'index et création des fichiers de couples clé-RID **FX1** et **FX2**;

[402X] - Si nécessaire, tri du fichier **FX2** par ordre de couples clé-RID croissants pour donner **FX'2**;

[403X] - Fusion virtuelle des fichiers **FX1** et **FX'2**;

[404X] - Rechargement du ou des fichiers contenant l'index avec les couples clé-RID ainsi ordonnés ou triés en respectant les paramètres d'espace libre.

Il est à remarquer que l'absence de l'opération de tri **[402]** et/ou **[402X]** résulte du fait que la lecture des lignes de l'espace de tables et/ou des pages d'index a pu être faite directement dans l'ordre, ce qui advient lorsque la base de données n'est pas trop désorganisée.

Ainsi l'utilisation de la lecture séquentielle qui permet de lire un grand nombre de pages en une seule opération (les unités de disques du système possédant une fonction de lecture en masse de blocs contigus) accélère de façon significative les opérations de réorganisation unitaires des bases de données indexées, notamment les bases de données indexées de type espace de tables.

Il est à noter pour terminer ce point de réorganisation unitaire que si les deux méthodes de réorganisation de la base de données **REORG/TABLE** et de réorganisation

de l'index **REORG/IND(ET)** peuvent être mises en œuvre conjointement, il apparaît qu'en pratique ces deux méthodes sont codées séparément et mises en œuvre comme des coroutines. Cette caractéristique est importante car elle donne plus de souplesse dans la préparation des plannings de réorganisation de l'ensemble des bases de données indexées notamment de type espace de table.

Après avoir présenté les grandes lignes de la réorganisation unitaire des bases de données indexées structurées en espaces de tables et de leurs index, il importe de revenir à la stratégie globale de la réorganisation de l'ensemble des bases de données du système d'information.

10 La gestion des réorganisations « en ligne »

La Stratégie

La stratégie globale intègre les réorganisations « en ligne » et « hors ligne ». Pour chaque objet du système informatique d'information on détermine le seuil de désorganisation à partir duquel une réorganisation « en ligne » est rentable. Si ce seuil est atteint sans qu'il est été possible de réorganiser l'objet « hors ligne » (manque de disponibilité dans la fenêtre « batch ») la réorganisation « en ligne » est lancée dès que l'activité de mise à jour est suffisamment faible pour que cette réorganisation soit valable. Cette réorganisation « en ligne » se fait en une seule opération sur une copie de l'objet en service. Les mises à jour éventuelles apparues pendant l'opération de réorganisation sont sauvegardées sur un fichier intermédiaire traité en fin de réorganisation (les pages modifiées étant repérées par l'heure de la modification), le processus étant répété pour les nouvelles mises à jour éventuelles jusqu'à épuisement. L'objet (la base de données) une fois réorganisé et mis à jour peut alors être remis en service.

En ce qui concerne les espaces de tables, c'est l'utilisation des parcours séquentiels d'index (en anglais index scans) par l'index primaire (index déterminant le critère de tri de la réorganisation) qui détermine la réorganisation des espaces de tables. En effet un espace de tables utilisé seulement en accès direct n'a pas besoin d'être réorganisé car les cartographies d'espace (en anglais space maps), seules utilisées pour déterminer la possibilité d'une insertion, sont suffisamment peu nombreuses pour être conservées dans le parc de mémoires tampons (en anglais buffer pool).

En ce qui concerne les index, l'accès direct (par pointeurs verticaux) à l'index est utilisé aussi bien en lecture qu'en mise à jour (c.à.d. modification du champ de données sauf indication contraire). Son temps d'Entrée/Sortie est déterminé par le fait que les pages représentant les deux derniers niveaux et qui dans la plus part du temps doivent être lues physiquement, sont ou non proches l'une de l'autre sur le disque.

C'est le temps de positionnement pour la lecture de la page « feuille » qui détermine la différence entre un index bien ou mal organisé.

Le parcours séquentiel (« index scan » utilisant des pointeurs horizontaux) ne représente qu'un coût marginal par rapport à la lecture concomitante de l'espace des tables, tout en ayant le même critère de performance que l'accès direct puisque si deux pages sont proches d'une troisième elles sont relativement proches l'une de l'autre. Le parcours séquentiel n'intervient donc pas dans la détermination de la nécessité d'une réorganisation.

Les Données utilisées

Il résulte de ce qui précède que les données à prendre en compte sont celles qui influent sur les performances des applications et sur le coût en termes de ressource de la réorganisation. Ces données sont de deux ordres ayant leurs origines soit dans les caractéristiques de l'objet à réorganiser, soit dans les caractéristiques des unités du système d'information (traitement et utilitaires):

i) les caractéristiques de l'objet à prendre en compte sont celles de déterminent le volume est l'état de l'objet :

S = le nombre de **CI** (bloc physique ou page) de 4 Ko de l'objet, index compris en cas d'espace de table,

NTL = le nombre total de lignes (RID si index), ou le nombre moyen par **CI**,

D = le taux de désorganisation, c'est à dire :

Pour un espace de tables :

DT = le pourcentage de lignes situées loin de la page qui contenait à l'origine la clé précédente au sens de l'index primaire ;

Pour un index :

DI = le pourcentage de pages feuilles situées loin de la page non feuille qui pointe dessus.

Pour calculer le nombre de lignes ou de RID on divise la taille utile du CI multiplié par le taux de remplissage estimé, par la taille typique de l'entrée.

· La taille utile d'un CI est voisine de 4 K pour un espace de tables et dépend de type de l'index et du nombre de sous pages et de la longueur des clés pour un index de type 1.

· Le taux de remplissage estimé de 0,85 pour un espace de tables et de 0,75 pour un index.

· La taille typique est la moyenne géométrique (racine nième du produit de n nombres) entre la taille maximum et la taille minimum ($t_{mg}^2 = t_{max} \cdot t_{min}$). Pour un index

multiple la taille du RID est le minimum, la taille d'une entrée unique le maximum. Pour un espace de tables ajouter 8 (préfixe + ID) à la taille de la ligne (à diviser par 3 en cas de compression).

ii) Les caractéristiques des unités à prendre en compte sont celles qui influent sur le temps d'Entrée/Sortie des traitements et utilitaires :

- le temps de positionnement moyen T_s ;
- le temps de latence T_l , c'est à dire le temps nécessaire pour faire un $\frac{1}{2}$ tour ;
- 5 - le nombre de blocs de 4K par piste de mémoire sur disque.

Est également à prendre en compte le profil applicatif (c'est à le nombre moyen d'un événement de l'application par unité de temps, par exemple l'heure).

Pour un espace de tables, l'information utilisée sera le nombre moyen de lignes lues par « index scan » (parcours séquentiel de l'index) par heure. En fait
10 remarquer qu'il n'est réellement utile de réorganiser un espace de tables que s'il y a des lectures en « index scans ».

Enfin on s'intéresse à la charge instantanée lorsque l'on cherche à déterminer l'opportunité de lancer une réorganisation « en ligne » à un instant donné. Jusqu'à présent on utilise le nombre moyen de mise à jour par objet et par heure. Il serait
15 opportun de travailler sur des prévisions à court terme de cette charge en utilisant des profils de journée d'exploitation.

Le Coût de la réorganisation :

Le coût d'une réorganisation en ligne, exprimé en temps d'Entrée/Sortie est donné par la formule **F1** suivante :

20 **$C = S.(c + d.D) / (1 - e.q.U)$** dans laquelle :

S est le nombre de pages ou taille de l'objet, index inclus ;

D est le taux de désorganisation ;

c,d,e sont des constantes de la machine (système informatique) ;

q = 1 si l'objet est un index,

25 sinon **q = nombre d'index + 1** ;

U est la moyenne horaire du nombre de mises à jour sur l'objet, les mises à jour étant pour l'essentiel des modifications de champs dans des lignes existantes (sans véritable effet de désorganisation).

On notera que la durée d'une réorganisation en ligne peut être très longue
30 (plusieurs heures pour une première phase de réorganisation en continu de durée t_0) entraînant la mise en mémoire provisoire, pendant la première phase de réorganisation, d'un premier paquet des mises à jour de l'objet. A la fin de cette première phase de durée t_0 (temps passé de réorganisation $T_r = t_0$) intervient une deuxième phase de la réorganisation plus courte ($t_1 < t_0$) que la première pour traiter
35 ces premières mises à jour, d'où l'apparition en fin de deuxième phase (avec $T_r = t_0 + t_1$) d'un second paquet de mises à jour plus petit que le premier, et ainsi de suite jusqu'à extinction des mises jour. Il en résulte que le temps total final de la réorganisation en

ligne Tf fait intervenir un facteur $k < 1$ fonction de U moyenne horaire du nombre de mises à jour sur l'objet.

Au départ de la réorganisation $Tr = 0$

Pour $Tr = t_0$, $C = Co = S.(c+d.D)$

5 Pour $Tr = t_0 + t_1$, $C = Co + k.Co = Co.(1+k)$ où $k < 1$

Pour $Tr = t_0 + t_1 + t_2$, $C = Co + k.Co + k.k.Co = Co.(1+k+k^2)$

Finalemment en fin de réorganisation $Tr = Tf$, avec:

$Tr = Tf = t_0 + \dots + t_n + \dots$, $C = Co + k.Co + k^2.Co + \dots + k^n.Co + \dots$

Soit un coût final $C = Co.(1+k+k^2+\dots+k^n+\dots) = Co/(1-k)$

10 où l'on retrouve la formule **F1** si l'on pose $k = e.q.U$

De la formule **F1** on tire le coût moyen horaire de réorganisation si t représente le nombre d'heures séparant deux réorganisations :

$Cm = [S.(c + d.D) / (1 - e.q.U)] / t$ soit **F1'**

Le Coût de ne pas réorganiser :

15 Le surcoût de ne pas réorganiser par heure de traitement vaut : **$C^* = Co^*.D$** , où **Co^*** est le surcoût horaire quand l'objet est totalement désorganisé.

D'une façon générale on considère que la désorganisation d'une base de données de type espace de tables résulte, quasi exclusivement des insertions de données dans la base (nouvelles lignes créées), soit :

20 **$D = i/S.r$** , où

i désigne le nombre d'insertions depuis la dernière réorganisation,

S désigne la taille de l'objet entier en pages,

r est le nombre moyen de lignes par page de l'entier,

Si l'on suppose le taux horaire d'insertion constant = **I**

25 **$D = I.t/S.r \Rightarrow C^* = Co^*.I.t/S.r$** où

I désigne le nombre d'insertions à l'heure, et

t le temps écoulé depuis la dernière réorganisation.

Le surcoût de ne pas avoir réorganisé depuis **t** heures est **$G = \text{Somme de } Co^*.I.t.dt / S.r = \frac{1}{2} . Co^*.I.t^2 / S.r$** soit

30 **$G = \frac{1}{2} Co^*.D.t$**

où **Somme de** désigne le signe mathématique de la fonction intégrale

Le surcoût moyen horaire de ne pas avoir réorganiser vaut donc : **$Cm^* = \frac{1}{2} . Co^*.D$** formule **F2**

Le moment optimal de la réorganisation :

35 Le coût moyen horaire total de la non réorganisation vaut : **$Cmt = Cm + Cm^*$**

Il importe de réorganiser lorsque **$Cmt = Cm + Cm^*$** est minimal.

Or

$$C_m = [S.(c + d.D) / (1 - e.q.U)] / t \text{ soit avec } S = l.t/D$$

$$C_m = l.(c + d.D) / (1 - e.q.U).r.D \text{ formule F1''}$$

5 $C_{mt} = C_m + C_m^* = \frac{1}{2}.Co^*.D + [l.(c + d.D) / (1 - e.q.U).r.D]$ soit la formule F3

Cette fonction du type $a.x + b + c/x$ passe par un minimum lorsque $a - (c/x^2) = 0$ soit $x^2 = c/a$, d'où pour la valeur de seuil D_s solution de l'équation suivante :

$$D_s^2 = 2 l.c / (1 - e.q.U) r Co^* \text{ formule F4}$$

10 Alternativement on montre que la fonction $C_{mt} = C_m + C_m^*$ admet un minimum lorsque les termes en D et en $1/D$ sont égaux ($a.x = c/x$), soit : $l.C / [(1 - e.q.U).r.D] = \frac{1}{2}.Co^*.D$

D'où on retrouve la formule F4.

Il y a lieu de réorganiser lorsque D^2 dépasse cette valeur D_s^2 , soit quand :

$$D^2 > 2 l.c / (1 - e.q.U) r Co^* \text{ ou, sous une autre forme, quand}$$

15 $R' = D^2 r Co^*. (1 - e.q.U) / 2.l.c > 1$

Ou quand $R' = R (1 - e.q.U) > 1$ avec

$$R = D^2 r Co^* / 2.l.c, \text{ soit la formule F5.}$$

Sachant que $e.q.U$ doit être le plus faible possible, la réorganisation en ligne interviendra pendant que U (moyenne horaire du nombre de mise à jour sur l'objet) est
20 faible à partir du moment où $R > 1$.

Ainsi le taux de seuil D_s est donné par la formule F5' suivante: $R = D_s^2 r Co^* / 2.l.c = 1$

i) Cas de l'Index

Si l'objet est un Index, chaque insertion coûte n lectures et 1 écriture si n est le
25 nombre de niveaux d'index et chaque relecture coûte n lectures. En supposant 3 niveaux d'index le temps d'accès correspondant aux deux derniers niveaux sera :

- $T_s + 2. T_l$, si l'index est réorganisé

- $2T_s + 2 T_l$, si l'index est désorganisé

T_s désignant le temps de déplacement de bras, T_l le temps de latence (1/2 tour du
30 bras). En cas d'index désorganisé le surcoût en temps d'accès par transaction est de T_s .

Si t^* est le taux de relecture de l'objet, c'est à dire le nombre moyen de fois qu'une entrée sera relue. Il est alors possible de connaître le nombre de transactions en fonction de l (nombre d'insertions à l'heure) soit : $(t^* + 1) l$.

Il en résulte que le surcoût horaire à ne pas réorganiser quand l'objet est
35 totalement désorganisé Co^* est $Co^* = (t^* + 1) l.T_s$ formule F6

On peut évaluer c (temps d'accès en E/S par la taille) en supposant dans le cas le plus défavorable que l'utilitaire de réorganisation « en ligne » lit une piste à la fois puis perd son positionnement, donc en négligeant les écritures.

$c = (T_s + 3T_l) / N_{bp}$ où N_{bp} est le nombre de pages de l'objet.

- 5 En posant $T_s / T_l = V$ qui est une caractéristique de l'unité de traitement on obtient

$R = D^2 r (t^* + 1) N_{bp} V / [2(V + 3)]$ formule F7

Ainsi le taux de seuil pour l'index D_{sl} est donné par la formule F7' suivante : $R = D_{sl}^2 r (t^* + 1) N_{bp} V / [2(V + 3)] = 1$

- 10 A titre d'exemple numérique non limitatif, on choisit :

$r = 200$ (clé de longueur 10 octets)

$N_{bp} : 12$ (3390)

$V = 1.76$ (3390-2)

$t^* = 1$ (une seule relecture)

- 15 On obtient $R = D^2 200.2.12.1.76 / 2.4.76 = D^2 8400 / 9.52$
soit $R = 887.D^2$

Il y aura lieu de réorganiser à partir de $R > 1$ (lorsque le taux U peut être considéré comme négligeable), soit pour $D^2 > 1/887$ ou $D > 0,0335$, soit à partir d'un taux de seuil de désorganisation de l'index $D_s = 3,35 \%$.

- 20 **ii) Cas de l'espace de tables**

Si l'objet est un espace de tables, une lecture séquentielle d'index (index scan) engendrera un surcoût de désorganisation par ligne déplacée de $T_s + T_l$.

Il en résulte un coût de désorganisation

$Co^* = L.l.(T_s + T_l)$

- 25 où L est la proportion des lignes accédées par index scan par rapport aux lignes créées.

Par ailleurs du fait que les index sont réorganisés simultanément, il faut tenir compte du coût associé aux accès directs à travers ces index,

soit $Co^* i = (t^* i + 1).l_i.T_s$

- 30 d'où le coût total moyen horaire à ne pas désorganiser

$Cm^* = \frac{1}{2}.L.l.D.(T_s + T_l) + \frac{1}{2} [S Co^* i.D_i^2 / D^2].D$

Soit

$Cm^* = \frac{1}{2}.L.l.D.(T_s + T_l) + \frac{1}{2} T_s. S l_i.D_i^2.(t^* i + 1)/D$

formule F9

- 35 Il est à noter :

- que $l_i = l$ dans le cas d'un espace de table monotable;

- que dans le cas général on peut supposer que $li = ki.l$ où ki désigne le cardinal relatif de la table associée à l'index i (et représentatif de la dimension relative de cette table). Cela revient à ignorer les mises à jour entraînant une modification de clé et à supposer que les tables ont des durées moyennes de vie de lignes comparables.

- 5 - que les t^*i représentent le taux d'accès direct au travers de chaque index donc que, pour une table donnée, $S t^*i$ représente le coût d'accès à cette table.

Supposons certaines réorganisations d'index faites, dans chaque intervalle ne contenant pas de réorganisation d'index, Di est une fonction linéaire du nombre d'insertions l donc de D .

10 $Di = ai.D - bi$, où $ai = li.S.r / (l.Si.ri) = ki / ki = 1$

d'où $Di = D - bi$

$$Cm^* = \frac{1}{2}.L.l.D.(Ts+Tl) + \frac{1}{2} Ts. S li.(D-bi)^2.(t^*i + 1)/D$$

$$Cm^* = \frac{1}{2}.L.l.D.(Ts+Tl) + \frac{1}{2} Ts. S li.D.(t^*i + 1)$$

$$- Ts. S li.bi.(t^*i + 1) + \frac{1}{2} S. li.bi^2.(t^*i + 1)/D$$

15 avec $Cm = l.(c + d.D) / (1 - e.q.U).r.D$ (coût moyen horaire pour réorganiser l'objet)

De façon analogue à ce qui a déjà été présenté ci-avant la fonction $Cmt = Cm + Cm^*$ admet un minimum local lorsque les termes en D et en $1/D$ sont égaux soit :

$$D^2.[L.l.(Ts+Tl)+Ts. S li.(t^*i + 1)]=2l.c/(1 - e.q.U).r + Ts. S li. D^2.(t^*i + 1)$$

20 $D^2.L.l.(Ts+Tl)+Ts. S li(D^2-bi^2)(t^*i + 1)=2l.c/(1-e.q.U).r$

En utilisant comme ci-devant $c = (Ts + 3.Tl) / Nbp$ et en substituant $bi = D -$

Di on obtient

$$R=r.Nbp[D^2.L (V+1) + V. S ki(2DDi-Di^2) (t^*i+1)]/ (2V+6)$$

Formule F12

25 Ou

$$R= r.Nbp[D^2.L (V+1) + V. S ki(D^2-bi^2) (t^*i+1)]/ (2V+6)$$

Formule 12'

Pour une base de données monotable on a $ki=1$, la formule devient

$$R= r.Nbp[D^2.L (V+1) + V. S (D^2-bi^2) (t^*i+1)]/ (2V+6)$$

30 Formule F13

Ainsi le taux de seuil DsT est donné par la formule F13' suivante: F13'

$$R=r.Nbp[DsT^2.L (V+1) + V. S (DsT^2-bi^2) (t^*i+1)] / (2V+6) = 1$$

A titre d'exemple numérique non limitatif, on choisit pour un espace de table monotable à 2 index:

35 $r=15$ (ligne de longueur 200 octets ; index 15% de la taille de l'espace de table)

$Nbp : 12$ (3390)

$$V = 1,76 (3390-2)$$

L = 5 (chaque ligne fait l'objet de 5 parcours séquentiels dans sa vie)

r1 = 200 (clé de longueur 10 octets)

r2 = 230 (clé de longueur 8 octets)

5 t*1 = 1 (une relecture directe)

t*2 = 2 (2 relectures directes)

En procédant comme présenté ci-avant, il apparaît que le premier index I1 (R1 = 887.DI1²) doit être réorganisé à partir de Dsl1 = 0,0355 = 3,35% et que le second index I2 (R2 = 1531.DI2²) doit être réorganisé à partir de Dsl2 = 0.0255 = 2,55%.

La base de données étant monotable on a ki = 1. La formule F12' se simplifie pour s'écrire :

$$R = (15.12)/(2.(4,76))[(D^2.5.(2,76)) + (1,76.2(D^2-b1^2)) + (1,76.3(D^2-b2^2))] \text{ soit}$$

15 $R = 180/9,52 [(13,8.D^2 + 3,52.D^2 + 5,28.D^2) - (3,52.b1^2 + 5,28.b2^2)]$

Soit $R = [427.DsT^2 - 66,5.b1^2 - 100.b2^2]/10.000 = 1$, formule F14, si les valeurs DsT, b1 et b2 sont directement exprimées en % (pourcent).

Il est clair de la formule F14 que la valeur de seuil de réorganisation de l'espace de table DsT sera supérieure à la limite Dinf solution de l'équation

20 $Dinf^2 = 10.000/427 = 23,4$, soit $Dinf = 4,8 \%$.

Après une réorganisation de l'espace de table et de ses deux index, tous les taux de désorganisation sont à zéro avec DT = DI1 = DI2 = 0.

La première réorganisation d'index interviendra pour Dsl2 = 2,55% (à un instant où la base de données et le premier index auront le même taux DT = DI1 = 2,55% (< Dinf), les taux étant proportionnels aux nombres d'insertion de nouvelles lignes (ou RID pour les index).

La seconde réorganisation d'index interviendra pour Dsl1 = 3,35% (à un instant où la base de données aura le même taux DT = 3,35% (< Dinf)).

En supposant pour simplifier que les index puissent être réorganisés dès que le taux ou seul limite est atteint, on peut examiner chaque intervalle défini entre deux réorganisations successives d'index et calculer par la formule F13 la valeur dans cet intervalle du taux limite pour l'espace de table (les valeurs de b1 et de b2 étant données par la valeur instantanée de la désorganisation de l'espace table Ds lors des deux dernières réorganisations d'index) et vérifier que cette valeur limite appartient bien à l'intervalle considéré ; à défaut l'intervalle considéré et la valeur limite calculée ne seront pas retenus.

On étudiera donc la suite d'intervalles suivante :

[0:Dsl2],[Dsl2:Dsl1],[Dsl1:2.Dsl2],[2.Dsl2:2.Dsl1], etc

où 2.Dsl2 et 2.Dsl1 correspondent respectivement à la seconde réorganisation du second et du premier index.

En introduisant les valeurs numériques la suite d'intervalles peut s'écrire :

5 **[0:2,55],[2,55:3,35],[3,35:5,1],[5,1:6,7], etc**

Compte tenu de la valeur de **Dinf=4,8%**, il y a lieu d'étudier l'éventualité d'une réorganisation de l'espace table dans l'intervalle **[3,35:5,1]**, avec **b1= 3,35** et **b2=2,55** (valeur de Ds lors de des deux premières réorganisation d'index – première de I1 et première de I2)

10 De la formule **F14**, on obtient **DsT²=11380/487 =26,5** et **DsT= 5,15%**

Cette valeur n'est pas acceptable car elle se situe en dehors de l'intervalle considéré. Il en résulte que la réorganisation de l'espace table ne pourra intervenir qu'après la seconde réorganisation du second index I pour laquelle **DsT** prendra la valeur **2x 2,55= 5,1**.

15 En reprenant cette procédure pour l'intervalle **[5,1:6,7]**, avec **b1= 3,35** et **b2= 5,1**, obtient **DsT²=13231/487= 31** et **DsT= 5,56%**.

Si on prend **U=0** la réorganisation de l'espace table et de ses deux index est souhaitable à partir d'un taux de désorganisation de **5,56%** survenant dans l'intervalle considéré à un instant précédant bien la seconde réorganisation du premier index.

20 Il est à noter en variante qu'il est intéressant d'appliquer également le critère de **rentabilité** présenté ci-avant à la réorganisation « hors ligne » (hl) dans la mesure où des données existent. Pour ce faire, il y a lieu de remplacer dans la formule **F1** le coefficient **c(el)** (temps d'Entré/Sortie en fonction de la taille de l'objet) par un coefficient correspondant à la réorganisation « hors ligne », par exemple du type **c(hl) =**
 25 **K.c(el)** où **K = Rapport du temps d'exécution de la réorganisation (hl) de l'objet sur le temps d'exécution de la réorganisation (el) du même objet dans des conditions moyennes d'exploitation de la base de données indexée de type espace de table correspondante.**

30 Il est ainsi possible de déterminer si une réorganisation « hors ligne » est ou non utile. Les priorités existantes basées sur les taux de désorganisation sont également remplacées par des priorités basées sur la rentabilité, un objet plus lu qu'un autre étant prioritaire pour la réorganisation à taux égal de désorganisation.

La gestion globale des réorganisations et copies « hors ligne » :

35 La réorganisation d'un fichier « hors ligne » est consommatrice de temps machine qu'il est en général souhaitable de minimiser pour réduire le coût total de l'opération.

Pour ce faire l'invention propose également une méthode de gestion ou d'ordonnancement préparée d'avance de l'exécution en temps réel de programmes utilitaires de réorganisation (et/ou d'images copie) de manière à utiliser au mieux les ressources du système informatique d'information dont la fenêtre « batch » (fenêtre de maintenance).

La réussite et l'efficacité d'une telle méthode repose sur la rapidité que ce soit la rapidité à préparer et à mettre à jour un plan d'activité ou planning des tâches à exécuter et/ou la rapidité à lancer l'exécution des programmes présélectionnés.

Avant d'exposer la méthode de gestion ou d'ordonnancement selon l'invention, il importe de présenter brièvement l'organisation et le fonctionnement d'un système informatique d'information comportant un ensemble de fichiers ou bases de données indexées répartis en divers objets réels constitués d'espace de tables, de partitions d'espaces de tables et d'index.

Le système d'information :

Les capacités « machine » du système d'information sont divisées en régions de traitement distinctes pilotées de façon unitaire par une région de contrôle. A chaque région de traitement sont affectées un certain nombre de tâches ordonnées et sont allouées les ressources nécessaires en mémoires virtuelles et périphériques (disques durs, imprimantes, mémoires sur supports externes etc...). Par souci d'efficacité en exécution et en allocation de ressources matérielles et/ou logicielles on affecte si possible à chaque région de traitement des tâches de même type, telles que par exemple pour des tâches généralement exécutables dans la fenêtre « batch » des réorganisations de fichiers ou bases de données (demandant beaucoup d'espace de disque) ou des copies images de fichiers (demandant des périphériques de sortie externe, lecteurs de bandes magnétiques, lecteurs/graveurs de CD-ROM, etc...).

A côté des réorganisations de fichiers tels que bases de données, la prise de copies des fichiers modifiés constitue un impératif de sécurité dans le domaine de l'exploitation de systèmes informatiques. On distingue généralement trois types de copie, la copie en continue ou copie « journal » (en anglais « log »), la copie incrémentale qui ne prend que les modifications apportées aux fichiers et la copie totale ou copie image.

En ce qui concerne les bases de données, les systèmes d'exploitation comportent assez souvent un dispositif logiciel de verrouillage interdisant la remise en service de la base de donnée avant la prise d'une copie image. Ceci crée une interaction entre les plannings de réorganisation et de copie.

Les plannings

Les plannings de réorganisation et de copie s'effectuent à partir de listes de priorités d'objets à réorganiser ou à copier qui seront traitées dans des régions de traitement différentes réservées soit à la réorganisation soit à la copie.

5 La méthode d'ordonnancement consiste à générer à l'avance les objets réels susceptibles de réorganisation au lieu de faire des combinaisons d'objets en temps réel. Le nombre d'objets en est plus grand pour les objets partitionnés, mais les priorités et par conséquent l'ordre logique d'enchaînement des réorganisations sont calculés initialement et rarement mis en cause.

10 De ce fait les objets à réorganiser peuvent être chaînés par priorité décroissante, même si certains sont mutuellement exclusifs.

La copie utilise d'autres régions et peut travailler sur les objets élémentaires. La liste des objets à copier est donc distincte de celle des objets à réorganiser avec un lien éventuel pour établir si un objet à copier est aussi à réorganiser ou fait partie d'un
15 objet à réorganiser.

Le processus d'ordonnancement ou de création de plannings peut se résumer de la façon suivante que se soit pour les copies ou les réorganisations :

- détermination ou présélection des objets à traiter en **priorité** par la réalisation d'un **planning rapide**.
- 20 - tentative de création d'une liste d'ordre d'exécution des traitements par placements en « retro-planning » des objets présélectionnés par priorités croissantes dans des régions de traitements disponibles, ce qui permet de s'assurer de l'heure maximum de démarrage d'un traitement. Ainsi le dernier objet sélectionné dans une région avec réussite pour compléter la liste devient le **prochain objet à traiter POC ou POR**. A
25 défaut une nouvelle tentative de création de liste sera effectuée sur d'autres hypothèses de placements ou d'affectations de régions de traitement.

Ce processus est renouvelé à chaque fin de traitement correspondant (copie ou réorganisation) en tenant compte de l'état réel des objets à cet instant.

L'intérêt de l'opération de « retro-planning » ou planning inversé est de
30 permettre à des gros objets non nécessairement prioritaires d'être traités avant qu'il ne soit trop tard (le temps de traitement encore disponible dans la fenêtre « batch » des régions de traitements libres devenant insuffisant).

Dans la réalité, les temps d'exécution sont variables et il y a lieu de prendre une marge de sécurité concernant les temps d'exécution ($K_c = 1,20$ pour les copies et
35 $K_r = 1,50$ pour la réorganisation).

En pratique chaque fois d'une région destinée à la réorganisation devient libre, l'algorithme de planning de réorganisation recherche l'objet réel dont la priorité

est maximale (planning rapide), le présélectionne et poursuit les tentatives d'affectations sur le même principe en examinant les conséquences sur sa sélection initiale (premier objet à réorganiser). En particulier :

- si l'objet en phase d'affectation est une partition contiguë à une partition d'un même espace de table déjà sélectionnée en premier, cette nouvelle partition sera ajoutée à la sélection initiale à titre d'objets liés dans la mesure où le nombre maximum de partitions sélectionnées pour être traitées ensemble n'a pas encore été atteint (toutes les partitions contiguës sélectionnées seront réorganisées en même temps) ;
- si l'objet en phase d'affectation est l'espace de table dont l'objet de la sélection initiale est un index, l'espace de table (et ses index) remplacera l'index de la sélection initiale et les éventuelles sélections des autres index seront annulés (les index d'un espace de table sont toujours réorganisés avec celui-ci).

Le processus sera partiellement réexécuté à chaque modification de l'objet sélectionné en premier (sélection initiale). De ce fait si de nombreux espaces tables du système d'information sont partitionnés, le nombre de remise en cause peut devenir important.

De façon analogue chaque fois d'une région destinée à la copie devient libre, l'algorithme de planning de copie recherche l'objet élémentaire dont la priorité est maximale (planning rapide), et refait un planning de réorganisation pour vérifier si cet objet est prévu ou non pour une réorganisation dans la fenêtre « batch ». Dans l'affirmative l'objet est enlevé provisoirement du planning des copies.

Par ailleurs, les priorités sont comparées entre les listes 'réorganisation' et les listes 'copie' ; des images copies plus prioritaires sont susceptibles d'empêcher une réorganisation du seul fait que celle-ci génère une copie et vice versa.

Le planning rapide

Le planning rapide consiste à déterminer les objets à traiter en limitant le temps global (la charge) à la somme des durées disponibles dans les régions de traitement.

Le planning rapide de la copie consiste à prendre les objets dans l'ordre décroissant des priorités il est donc aussi simple à réaliser qu'un tri d'objet. Il permet de donner une évaluation rapide du temps de fenêtre résiduel en fonction de la priorité. Il peut être ajusté en fonction des objets ajoutés par le planning rapide de réorganisation.

Le planning rapide de réorganisation prend en compte les liens entre objets réels :

- si un objet est contenu dans un autre, on planifie le contenant ;

· si deux objets ont une partition commune sans que l'un contienne l'autre, on planifie celui qui à la plus haute priorité.

Le planning rapide de réorganisation travaille également dans l'ordre décroissant mais il est capable de remettre en question l'élimination de gros objets en fin de fenêtres.

La rapidité des plannings rapides tient à ce qu'ils sont limités dans leurs retours en arrière et ne remettent pas en cause les priorités des objets.

La boucle la plus intérieure est celle qui cumule les temps de copie en testant si un objet est à la fois copié et réorganisé. Il est possible d'introduire un curseur sur l'objet à copier, sachant que le parcours complet n'est nécessaire que si des incompatibilités ou remise en causes ont été détectées.

Les priorités

Les priorités relatives de copies et réorganisation sont prises en compte de manière suivante :

Pour les réorganisations nécessitant une image copie, on vérifie qu'elle est possible (planning rapide) en tentant d'insérer l'objet (s'il n'y est pas déjà) au planning provisoire de copie avec sa priorité de réorganisation.

Ainsi si ce sont les régions de copie qui limitent le nombre d'objets, cette limite se fera à des priorités voisines tant pour la réorganisation des objets nécessitant une image copie que pour la copie des autres objets.

Par contre si ce sont les régions de réorganisation qui limitent le nombre d'objets, des copies d'objets beaucoup moins prioritaires peuvent avoir lieu.

Les spécifications de la méthode de gestion et d'ordonnancement de l'exécution de logiciels utilitaires de réorganisation de bases de données indexées.

Les Données

Les objets élémentaires

Les objets élémentaires sont les plus petits objets susceptibles d'un traitement (réorganisation et/ou copie) et pour lesquels on est capable de mesurer un taux de désorganisation.

Ces objets sont utilisés pour l'image copie et comme intermédiaires pour générer des objets de réorganisation.

Les objets élémentaires pour lesquels une image copie est à faire sont chaînés (liste PRIOCOPIE) entre eux par priorités décroissantes et tailles croissantes (la priorité la plus haute étant donnée à l'objet dont la dernière copie remonte le plus loin dans le temps). Les ajouts après réorganisation se font en effet en tête de la liste PRIOCOPIE.

Tous les objets élémentaires faisant l'objet d'une copie (y compris, ultérieurement, après leur réorganisation) comportent un pointeur de pile dans l'ordre des priorités croissantes qui est valorisé par le planning de copie.

Les objets réels

- 5 Les objets réels sont ceux qui peuvent être réorganisés sans entraîner la réorganisation d'autres objets (ex ; espaces de tables + index non pris isolément). Les objets réels sont créés à partir d'objets élémentaires (espaces de tables, index et partitions).

- 10 La **priorité** de chaque objet réel est le barycentre des taux de désorganisation des composants avec leurs tailles pour coefficients. On rappelle que le taux de désorganisation d'un objet est donné par le rapport nombre d'insertions (lignes ou RID) sur la taille de l'objet. Les objets réels sont rangés dans la liste des priorités de réorganisation (liste **PRIOREORG**) par priorités décroissantes.

- 15 La priorité étant définie comme le rapport **gain / coûts**, le gain d'un objet est calculé comme le produit de sa priorité par le temps de sa réorganisation.

Les objets réels sont chaînés avant-arrière afin de permettre le parcours dans les deux sens et les opérations de suppression et de changement de priorité qui ont lieu chaque fois qu'une réorganisation est lancée.

- 20 Le planning rapide utilise également une liste de suppression (liste **SUPP**) permettant de différer la suppression d'objets jusqu'à la confirmation de la nécessité de ces suppressions par la sélection de l'objet courant incompatible avec les objets supprimés.

- 25 Il est possible d'introduire une sous-chaîne d'objets sélectionnés (ou ayant fait l'objet de tentatives de sélection) afin d'éviter, lors de la phase de remise en cause d'avoir à examiner de nombreuses combinaisons de partitions incompatibles avec les choix antérieurs.

L'objet partitionné

- 30 Pour un espace de table contenant n partitions il y a $2^n - 1$ objets réels possibles pour la réorganisation, ce qui est irréaliste. On se limite aux combinaisons de partitions contiguës dans l'ordre des priorités soient à $\frac{1}{2} n.(n+1)$ objets. Si le nombre maximum de partitions réorganisées simultanément est $p < n$, le nombre d'objets réels se réduit à $\frac{1}{2} p.(2n-p+1)$.

Par ailleurs il est nécessaire de pouvoir répondre rapidement aux questions suivantes :

- 35 i) - un objet est-il contenu dans un autre ?
ii) - deux objets ont-ils une partition commune ?

Pour faciliter la réponse à ces questions on associe à chaque objet un tableau de bits (en anglais **bitmap**) marqué à **1** pour chaque partition de l'objet et on procède à des comparaisons des « bitmaps ».

Objets globaux

- 5 Les objets globaux sont les espaces de tables. Il sont destinés à contenir les informations relatives à la **partie planifiée** :

Les informations temporaires utilisées par le planning rapide sont :

- un pointeur vers une liste de ses sous-objets réels sélectionnés, une valeur non nulle indiquant qu'une partie de l'objet est à réorganiser ;
- 10 - un indicateur de réorganisation de l'ensemble de l'espace de tables ;
- un indicateur de copie d'une partie de l'espace de table ;
- une « bitmap » des réorganisations des partitions de l'espace de tables .
- une « bitmap » des réorganisations des partitions d'index.

- 15 Il y a un pointeur de chaque objet élémentaire ou réel sur l'objet global correspondant.

Enfin les objets réels sélectionnés d'un même espace de table sont chaînés entre eux par le planning rapide dans la liste de sélection (liste SELECT/OR) de l'objet global.

Régions de traitement

- 20 Chaque région de traitement est affectée soit à la copie soit à la réorganisation. La recherche du prochain objet à réorganiser ou copier a lieu quand une région est libérée par la fin de la réorganisation ou la copie précédente.

Le planning consiste donc à valoriser un pointeur de la région de traitement vers le prochain objet à traiter.

- 25 Afin de permettre l'anticipation, on conserve pour chaque région l'heure relative estimée de fin du traitement en cours dans la région, laquelle est initialement nulle.

Les Traitements :

- 30 Les traitements ci-après sont données à **titre d'exemple non limitatif** d'un mode d'exploitation de la mise en œuvre de la méthode de gestions des réorganisations et des copies d'un ensemble de bases de données indexées selon l'invention. Sans sortir du cadre de l'invention, ces traitements peuvent faire l'objet de variantes et de modifications diverses pour remplir des fonctions techniquement identiques ou équivalentes .

35 **A Planning rapide**

On rappelle que le planning rapide a pour seul but de déterminer les objets pouvant être réorganisés dans la durée résiduelle de la fenêtre « batch » et non de

calculer l'heure de début des réorganisations. Les seules informations utilisables sont l'indicateur de sélection de chaque objet réel et les indicateurs d'affectation (partie planifiée et pointeurs).

Le **planning rapide de réorganisation PRR*** comporte les six phases
5 opérationnelles suivantes :

Initialisation

Boucle sur objets

Elimination des intersections d'objets

Remise en cause des choix précédents

10 **Vérification de la suffisance du temps copie**

Sélection de l'objet courant

La phase opérationnelle d'Initialisation comporte les opérations suivantes :

- Pour tout objet réel de la liste, **annuler** l'indicateur de sélection **IndS/OR** ;
- Pour tout objet global **OG** de la liste, **annuler** le pointeur de liste de sélection
- 15 **PointLS/OG** et **annuler** la partie planifiée de l'objet **PPO/OG**;
- **Initialiser** le limiteur de remises en cause à 10.000 ;
- **Initialiser E** = temps minimum de traitement = minimum de (durée unitaire de réorganisation + copie des objets réels à réorganiser **OR**) ;
- **Initialiser Temps résiduel de copie TRC** = temps restant après exécution des copies
- 20 en cours et des copies liées aux réorganisations en cours ;
- **Initialiser Temps résiduel de réorganisation TRR** = temps restant après exécution des réorganisations ;
- **Initialiser** : Prochain Objet à Copier **POC** = premier objet de la liste finale (en réalité une liste virtuelle car seul le premier objet de cette liste importe).

25 **Nota :**

1) Cette phase d'Initialisation débutant le processus de **PRR*** est notamment lancée à chaque fin d'une tâche de réorganisation (réorg + copie) dans une région de traitement correspondante.

De la même façon à chaque fin de tâche de réorganisation les listes de
30 priorités **PRIOREORG** pour les objets réels et **PRIOCOPIE** pour les objets élémentaires sont remises à jour.

2) Au tout départ des réorganisations on a :

TRC = Durée fenêtre « batch » x Nb de régions de copie, TRR = Durée fenêtre « batch » x Nb de régions de réorganisation.

35 **La phase opérationnelle de Boucle sur objets comporte les opérations suivantes :**

Pour tout objet réel de la liste **PRIOREORG** jusqu'à ce que **TRR < E** :

- **Marquer** la sélection possible
- **Annuler** le pointeur de la liste **SUPP**
- **Initialiser** Temps Provisoire de Copie **TPC = TRC**
- **Initialiser** Temps de Réorganisation **TR = Durée de réorganisation de l'objet courant OR**

5 **La phase opérationnelle d'Elimination des intersections d'objets comporte les opérations suivantes :**

 Pour tous les objets globaux associés à l'objet courant **OR**, **Faire :**

Si le pointeur de la liste de sélection **SELECT/OR** de l'objet global associé à l'objet courant est non nul (ce qui indique qu'il y a déjà au moins un objet réel sélectionné),

10 **Faire :** Si la partie planifiée de l'objet global **OG** est contenue dans la partie planifiée de l'objet courant **OR :**

Faire pour tout objet réel **B** de la liste de sélection :

 Si **B** marqué sélectionné alors :

- i) **Ajouter** le temps additionnel de copie de **B** au **TPC**
- 15 ii) **Soustraire** la durée de réorganisation de **B** du temps de réorganisation **TR**
- iii) **Placer B** dans la liste de suppression **SUPP**

sinon

 Si la partie planifiée de l'objet global **OG** a une intersection non vide avec la partie planifiée de l'objet courant :

20 **Marquer** sélection de l'objet courant **OR impossible**

La phase opérationnelle de Remise en cause des choix précédents comporte les opérations suivantes :

 Si sélection possible

25 Si durée de réorganisation + copie de l'objet courant **OR** > durée résiduelle dans la fenêtre « batch » de dans la région disponible, alors

Marquer sélection de l'objet courant **impossible**

Sinon (durée réorg + copie < durée résiduelle « batch »)

 Si le temps de réorganisation (durée de réorganisation de l'objet courant) > durée résiduelle de réorganisation

30 **Initialiser** : gain résiduel = gain de l'objet courant

Initialiser : temps restant à gagner = durée de réorganisation de l'objet courant – durée résiduelle de réorganisation

35 **Faire** pour tout objet réel **B** (déjà sélectionné) de la liste **SELECT/OR** en commençant par le précédent de l'objet courant **OR** et dans l'ordre de priorités croissantes tant que le limiteur des remises en cause >0 et que le temps restant à gagner ≤ 0 :

Si **B** est sélectionné, sans être sous-objet de l'objet courant, et si **gain B < gain résiduel (= gain de l'objet courant OR)**

Soustraire le gain de **B** du gain résiduel

Soustraire la durée de réorganisation de **B** du temps restant à gagner

5 **Ajouter** le temps additionnel de copie de **B** au temps provisoire de copie

Placer B dans la liste de suppression **SUPP**

Décrémenter le limiteur de remises en cause

Si le temps à gagner est >0

Marquer sélection de l'objet courant **impossible**

10 Sinon sélection possible soit parce que le temps gagner après l'élimination de **B** est ≤ 0 , soit par ce que temps de réorganisation (durée de réorganisation de l'objet courant) < durée résiduelle de réorganisation.

La phase opérationnelle de Vérification de la suffisance du temps copie comporte les opérations suivantes :

15 Si sélection possible

Initialiser : Temps additionnel de copie de l'objet courant = temps de copie de l'objet courant

Si copie à faire pour l'objet courant **OR**,

20 **Faire** pour tout objet de la liste des objets restant à copier (liste **PRIOCOPIE**) tant que la priorité de l'objet à copier > à la priorité de l'objet à réorganiser

Soustraire le temps de copie de l'objet à copier du temps provisoire de copie

Mais si l'objet à copier est un sous-objet de l'objet courant à réorganiser

25 **Soustraire** le temps de copie de l'objet à copier du temps provisoire de copie (= temps résiduel de copie)

Si le temps additionnel de copie de l'objet courant > temps provisoire de copie

Marquer sélection de l'objet courant **impossible**

30 **La phase opérationnelle de Sélection de l'objet courant comporte les opérations suivantes :**

Si sélection possible

Positionner l'indicateur de sélection **IndS** de l'objet courant ;

Positionner l'indicateur d'affectation de l'objet global associé à l'objet courant ;

Noter la partie planifiée **PPO** de l'objet global associé à l'objet courant **OR**;

35 **Valoriser** : pointeur (suite) de sélection de l'objet courant **OR** = pointeur liste de sélection **PointLS** de l'objet global **OG** ;

Valoriser : pointeur liste de sélection **PointLS** de l'objet global = adresse de l'objet courant **OR** (l'objet courant est ajouté la liste de sélection **SELECT/OR**);

Soustraire le temps additionnel de copie de l'objet courant du temps résiduel de copie **TRC** ;

5 **Soustraire** la durée de réorganisation de l'objet courant **OR** du temps résiduel de réorganisation **TRR** ;

Si le pointeur de liste de suppression est non nul

Faire pour tout objet réel B de la liste de suppression **SUPP**:

Si B est marqué sélectionné

10 **Annuler** l'indicateur de sélection de B

Ajouter la durée de réorganisation de B au temps résiduel de réorganisation **TRR**.

15 **Nota** : il n'est pas utile d'enlever B de la liste de sélection car seuls les objets réels **OR** définitivement sélectionnés et **marqués par l'indicateur de sélection** seront considérés comme tels et examinés par la suite.

B Prochain Objet à Copier

20 La recherche du prochain objet à copier est effectuée chaque fois qu'une région est libérée par une image copie terminée ou qu'une région est disponible et qu'un nouvel objet a été introduit suite à une réorganisation. Les objets doivent être triés par priorités décroissantes et tailles croissantes, les objets provenant des réorganisations étant placés en tête de liste.

25 Le planning s'effectue du moins prioritaire au plus prioritaire et il est possible qu'un gros objet (longue copie) déborde de la région de copie qui lui a été attribuée. Cette situation est acceptable dans la mesure où l'on s'est assuré que l'objet tenait dans la région du point de vue Temps Résiduel de Copie et du fait qu'une technique valable de planification consisterait à enlever de cette région un objet moins prioritaire déjà sélectionné et de le placer dans une autre région, ce qui ne changerait pas la sélection pour cette région.

30 Le pointeur prochain objet à copier est valorisé pour chaque région, mais pour la raison citée précédemment, il est recommandé de refaire le planning à chaque fois, même initialement quand toutes les régions sont disponibles.

Enfin il est possible que le planning copie choisisse un objet qui n'appartienne pas (encore) à la liste de copie parce sa réorganisation n'est pas terminée. Dans ce cas on se met en attente (région disponible).

35 **Le traitement Prochain Objet à Copier comporte les deux phases opérationnelles suivantes :**

Sélection des objets à copier

Rétroplanning de copie**La phase opérationnelle de Sélection des objets à copier**

comporte les opérations suivantes :

Faire un planning rapide de réorganisation PPR*;

5 **Initialiser** les piles de sélection de copie à vide ;

Initialiser : temps résiduel de copie **TRC** = temps total de fenêtre « batch » restant disponible dans les régions « copie » après exécution des copies en cours ;

Initialiser EC = minimum (durée unitaire de copie des objets à réorganiser **OR**)

Faire pour toutes les régions de réorganisation,

10 **Si** région active

Si image copie de l'objet en cours de réorganisation à faire,

Faire pour tout sous-objet pouvant être copié séparément (partition)

Soustraire le temps de copie du sous-objet de la durée résiduelle

de copie

15 **Placer** le sous-objet courant en haut de la première pile de sélection de copie.

Faire pour tout objet réel sélectionné de la liste des objets **OR** à réorganiser

SELECT/OR dans l'ordre des priorités décroissantes :

Si image copie de l'objet courant **OR** à faire,

20 **Faire** pour tout sous-objet pouvant être copié séparément (partition)

Soustraire le temps de copie du sous-objet de la durée résiduelle de copie

Placer le sous-objet courant en haut de la première pile de sélection

de copie.

25 **Faire** pour tout objet de la liste des objets à copier **OC** dans l'ordre des priorités décroissantes, tant que la durée résiduelle de copie **TRC** > **EC**

Si le temps de copie de l'objet **OC** \leq durée résiduelle de copie **TRC**,

Si l'objet à copier **OC** n'est pas déjà dans la première pile (l'objet **OC** est un sous-objet ; vérification possible par test de la partie réorganisée de l'objet global en cas de partition),

30 **Soustraire** le temps de copie du sous-objet de la durée résiduelle de copie

Placer le sous-objet courant en haut de la seconde pile de sélection

de copie.

Placer la première pile de sélection de copie sur la seconde (constituant ainsi la liste **SELECT/COPIE**).

35 **La phase opérationnelle de Rétroplanning de copie comporte les opérations suivantes :**

Faire pour toute région de copie :

Durée de la région = Durée de la fenêtre – heure relative estimée de fin de traitement en cours dans la région

Durée consommée région = 0

Initialiser le premier pointeur objet à copier de la région courante à 0.

5 Faire de haut en bas pour tout objet de la pile de sélection de copie :

Si temps de copie de l'objet courant **OC** \leq durée de la région la plus longue ;

Rechercher une région dont la durée consommée est minimale ;

Si temps de copie de l'objet **OC** $>$ durée de la région courante – durée consommée

10 **Rechercher** une région pour laquelle : durée de la région – durée consommée de la région – temps de copie de l'objet courant soit minimal et positive ou nulle,

Si aucune région ne convient

Rechercher une région pour laquelle : durée de la région – durée consommée de la région – temps de copie de l'objet courant soit maximale (c.à.d. minimale en valeur absolue) et le temps de copie de l'objet courant \leq durée de la région,

Si l'objet courant **OC** appartient à la seconde pile,

Valoriser :pointeur premier objet à copier **POC** de la région courante = adresse de l'objet courant **OC**,

20 **Ajouter** la durée de copie de **OC** à la durée consommée de la région courante.

C Prochain Objet à Réorganiser

La recherche du prochain objet à réorganiser est effectuée chaque de fois qu'une région est libérée par une réorganisation terminée. Préalablement les priorités ont dû être calculées ou corrigées et les objets réels triés dans l'ordre des priorités décroissantes.

25 Comme pour la copie, le planning s'effectue du moins prioritaire au plus prioritaire. On commence par planifier les copies engendrées par les réorganisations, ce qui permet de calculer une heure limite de fin de réorganisation pour chaque objet devant être copié.

30 Le temps disponible de chaque région de réorganisation est divisé en trois parties :

1 - la durée consommée effectivement par les réorganisations planifiées

2 - la durée gaspillée par le temps réservé à la copie après réorganisation

3 – la durée résiduelle allant du début à la première réorganisation

35 Le processus tente de remplir la durée gaspillée par des réorganisations d'objets ne nécessitant pas de copie.

Quand un objet à réorganiser est sélectionné, il est retiré de la liste ainsi que ses index (si applicable) et tous les autres objets qui ont des partitions communes avec cet objet. Les priorités des autres objets pouvant partager un index avec celui-ci sont également recalculées (pour tenir compte de l'impact de la réorganisation de l'index) et ces objets sont replacés aux bons endroits dans la liste des objets réels à réorganiser.

Le traitement Prochain Objet à Réorganiser comporte les quatre phases opérationnelles suivantes :

Rétroplanning de copie

10 **Rétroplanning de réorganisation**

Identification de la région de traitement

Inscription de l'objet au planning

La phase opérationnelle de Rétroplanning de copie comporte les opérations suivantes :

15 **Faire un planning rapide**

Faire pour toute région de copie :

Durée consommée de la région = 0

Faire pour tout objet réel OR sélectionné de la liste des réorganisations SELECT/OR (dans l'ordre des priorités croissantes)

20 **Initialiser : 'temps à réserver pour copie de l'objet courant ' = 0**

Si image copie de l'objet courant **OR** à faire,

Faire pour tout sous-objet pouvant être copié séparément (partition),

Rechercher une région dont la durée consommée est minimale, ci-après région courante

25 **Ajouter le temps de copie du sous-objet à la durée consommée de la région courante,**

Valoriser 'temps à réserver pour copie de l'objet courant' = max (temps à réserver pour copie de l'objet courant OR, durée consommée de la région courante)

30 **La phase opérationnelle de Rétroplanning de réorganisation comporte les opérations suivantes :**

Faire pour toute région de réorganisation :

Durée de la région = Durée de la fenêtre – heure relative estimée de fin du traitement en cours dans la région.

35 **Durée consommée dans la région = 0**

Durée gaspillée dans la région = 0

Pointeur premier objet à réorganiser de la région courante = 0

Faire pour tout objet réel **OR** sélectionné de la liste **SELECT/OR** dans l'ordre des priorités croissantes :

Faire pour chaque région

Si l'image copie de l'objet courant **OR** à faire,

- 5 **Calculer** dernier délai de la région = durée de réorganisation de l'objet courant + **max** de [temps à réserver pour copie de l'objet courant, (durée consommée de la région + durée gaspillée de la région)]

Sinon,

- 10 Calculer le dernier délai de la région = durée consommée de la région + **max** de [(durée de réorganisation de l'objet courant, durée gaspillée de la région)]

La phase opérationnelle de Identification de la région de traitement comporte les opérations suivantes :

Rechercher une région dont le délai est minimal,

Si dernier délai de la région courante > durée de la région courante,

- 15 **Rechercher** une région pour laquelle : [durée de la région courante - dernier délai de la région courante] est minimale et positive ou nulle,

Sinon, si aucune région ne convient,

- 20 **Rechercher** une région telle que : [valeur absolue de (durée de région - dernier délai de la région courante)] est minimale, et [durée de réorganisation + copie objet courant] \leq durée de la région.

La phase opérationnelle de Inscription de l'objet au planning comporte les opérations suivantes :

Si image copie de l'objet courant **OR** à faire ou durée gaspillée de la région courante = 0,

- 25 **Valoriser: pointeur premier objet à réorganiser POR de la région courante = adresse de l'objet courant OR,**

Ajouter la durée de réorganisation de l'objet courant **OR** à la durée consommée de la région courante,

Valoriser : durée gaspillée de la région courante = dernier délai de la région courante - durée consommée de la région courante.

- 30 Lorsque l'une des deux régions de traitement identifiées concernées (réorganisation ou copie) devient disponible, le système lancera, en l'absence de tout événement de remise à zéro **RAZR=1** ou **RAZC=1** survenu entre temps pour le traitement concerné, soit une réorganisation de l'objet **POR** (prochain objet à réorganiser) correspondant, soit une copie de l'objet **POC** correspondant (prochain objet à copier).
- 35

Une fois le traitement lancé le processus d'identification **IDPOR** et/ou **IDPOC** reprendra, les listes **PRIOREORG** et **PRIOCOPIE** étant mises à jour en continue,

jusqu'à épuisement du temps de réorganisation et/ou de copie imparti aux diverses régions de réorganisation et de copie correspondantes dans la fenêtre « Batch ».

5 Bien entendu la méthode de gestion des réorganisations et de copies dans un ensembles de bases de données indexées d'un système d'information selon l'invention n'est pas limitée au mode de mise en œuvre décrit ci-dessus et couvre d'autres variantes et modifications qui pourraient être appréciées par l'homme de métier tout en restant dans le cadre des revendications présentées ci-après.

10 En particulier la méthode de gestion selon l'invention utilisée pour à la réorganisation « hors ligne » (off line) des bases de données est également applicable à la réorganisation des bases de données « en temps réel » ou « en ligne » (on line) selon une méthode mixte, la réorganisation « hors ligne » et la réorganisation « en ligne » n'étant pas exclusive l'une de l'autre.

REVENDEICATIONS

1. Méthode de gestion des réorganisations dans un ensemble de bases de données indexées d'un système informatique d'information adaptée à la réorganisation « hors ligne » dans au moins une région de traitement de réorganisation du système, caractérisée en ce qu'elle comporte les phases opérationnelles suivantes :

(10) - Création et maintien par mise à jour en continu d'une liste **PRIOREORG** des objets à réorganiser par priorité décroissante fonction de l'état de désorganisation des objets à réorganiser ;

(20) - Exécution du processus **IDPOR** d'identification du premier ou prochain objet à réorganiser **POR** avec, à chaque événement **RAZR=1** interne au système pour lequel la sélection du premier ou prochain objet à réorganiser doit être remise en cause, interruption et remise à zéro et nouvelle exécution du processus **IDPOR**, lequel processus comportant les étapes suivantes :

(21) - Etablissement d'un planning rapide de réorganisation **PRR** à l'aide de la liste **PRIOREORG** et du temps opérationnel résiduel disponible **TRR** dans l'ensemble des régions de réorganisation et d'une liste de sélection **SELECT/OR** d'objets à réorganiser **OR** ordonnés par priorités décroissantes optimisées en fonction du gain de la réorganisation de chaque objet **OR**, ledit gain étant défini comme le produit d'un facteur représentatif du taux de désorganisation d'un objet **OR** par le temps de réorganisation de cet objet;

(22) - Etablissement d'un rétroplanning de réorganisation dans lequel pour tout objet courant à réorganiser **OR** extrait de la liste **SELECT/OR** par priorité croissante pour favoriser le traitement avancé des objets de plus grandes tailles possibles, on calcule pour chaque région de réorganisation le dernier délai de la région **DDRR**, représentant dans la fenêtre « batch » allouée à la réorganisation « hors ligne » le temps minimal nécessaire à la réorganisation de l'objet courant **OR**, égal dans le cas présent à la durée de temps consommé **DCRR** plus le temps de réorganisation de l'objet courant **DROR**;

(23) - Identification de la région de traitement **RTR** de l'objet courant **OR** par adéquation optimisée entre la durée de traitement disponible dans ladite région **DRR** avec la durée de traitement **DROR** nécessaire pour la réorganisation de l'objet **POR** avec **DRR - DDRR minimal ≤ 0** ;

(24) - Inscription de l'objet courant **OR** comme prochain objet à réorganiser **POR** au planning de la région identifiée en valorisant le pointeur **POR** correspondant à l'adresse de l'objet courant **OR** et en augmentant la durée de temps consommé **DCRR** dans la région du temps de la réorganisation **DROR**.

(30) - Lancement de la réorganisation de l'objet **POR** dès que la région de traitement de réorganisation **RTR** identifiée est libérée en l'absence de tout événement **RAZR=1** survenu entre temps ; Remise à zéro et nouvelle exécution du processus **IDPOR**.

5 2. Méthode de gestion selon la revendication 1, caractérisée en ce que l'ordre optimal des réorganisations est coordonné avec l'ordre des copies de sauvegarde d'objets du système d'information en tenant compte des réorganisations en cours ou juste effectuées.

10 3. Méthode de gestion selon la revendication 2, caractérisée en ce que l'ordre des copies de sauvegarde est modifié par l'exécution en première priorité, pour au moins un objet à réorganiser **OR**, de la copie dudit objet **OR** au plus tôt en fin de traitement de réorganisation.

15 4. Méthode de gestion des réorganisations d'un ensemble de bases de données indexées d'un système d'information adaptée à la réorganisation « hors ligne » dans au moins une région de traitement de réorganisation selon l'une des revendications 1 à 3, caractérisée en ce qu'elle intègre l'organisation des copies dans au moins une région de traitement de copie et comporte les phases opérationnelles suivantes:

(40) - Création et maintien par mise à jour en continu :

20 - d'une liste **PRIOREORG** des objets à réorganiser par priorité décroissante fonction de l'état de désorganisation des objets à réorganiser ; et
- d'une liste **PRIOCOPIE** des objets à copier par priorité décroissante fonction de l'ancienneté des dernières copies des objets à copier.

(50) – Surveillance en continu de l'apparition dans le système d'information de tout événement :

25 - **RAZR=1** pour lequel la sélection du premier ou prochain objet à réorganiser **POR** doit être remise en cause, notamment une fin d'une tâche de réorganisation dans une région de traitement de réorganisation et/ou la libération d'une zone de traitement pour la réorganisation ou la copie ; et/ou

30 - **RAZC=1** pour lequel la sélection du premier ou prochain objet à copier **POC** doit être remise en cause, notamment une fin d'une tâche de copie dans une région de traitement de copie, une copie prioritaire en fin de réorganisation et/ou la libération d'une zone de traitement pour la réorganisation ou la copie, avec

pour tout événement **RAZR=1**, lancement de l'exécution du processus **IDPOR** d'identification du premier objet à réorganiser **POR** ou, si le processus **IDPOR** est en cours d'exécution, interruption et nouveau lancement du processus **IDPOR** ; et/ou

35 pour tout événement **RAZC=1**, lancement de l'exécution du processus **IDPOC** d'identification du premier objet à copier **POC** ou, si le processus **IDPOC** est en cours d'exécution, interruption et nouveau lancement du processus **IDPOC** ;

- le processus **IDPOR** comportant les opérations suivantes :

(51) - Etablissement d'un planning rapide de réorganisation **PRR*** à l'aide de la liste **PRIOREORG** et du temps opérationnel de réorganisation résiduel disponible **TRR** dans l'ensemble des régions de réorganisation et d'une liste de sélection **SELECT/OR** d'objets à réorganiser **OR** ordonnés par priorités décroissantes optimisées en fonction du gain de la réorganisation de chaque objet **OR**, ledit gain étant défini comme le produit d'un facteur représentatif du taux de désorganisation d'un objet **OR** par le temps de réorganisation de cet objet;

(52) - Etablissement d'un planning rapide de copie **PRC** à partir de la liste **SELECT/OR**, puis de la liste **PRIOCOPIE** et dans la limite du temps opérationnel de copie résiduel disponible **TRC** dans l'ensemble des régions de copie, d'une liste de sélection **SELECT/OC** en mémoire de type LIFO ou pile pour d'objets à copier **OC** empilés par priorités décroissantes;

(53) - Etablissement d'un rétroplanning de copie dans lequel pour tout objet courant à réorganiser extrait de la liste **SELECT/OR** par priorité croissante, on calcule pour chaque région de copie le temps à réserver **TCR** pour la copie de l'objet **OR** après réorganisation ;

(54) - Etablissement d'un rétroplanning de réorganisation dans lequel pour tout objet courant à réorganiser **OR** extrait de la liste **SELECT/OR** par priorité croissante pour favoriser le traitement avancé des objets de plus grandes tailles possibles, on calcule pour chaque région de réorganisation le dernier délai de la région **DDR**, représentant dans la fenêtre « batch » allouée à la réorganisation « hors ligne » le délai minimal nécessaire à la réorganisation de l'objet **OR**, selon une formulation distincte si une copie de l'objet **OR** en fin de réorganisation doit être faite ou non;

(55) - Identification de la région de traitement **RTR** de l'objet courant **OR** par adéquation optimisée entre la durée de traitement disponible dans ladite région **DRR** avec la durée de traitement **DROR** nécessaire pour la réorganisation de l'objet courant **OR** avec $DRR - DDL_{\text{minimal}} \leq 0$ sans copie à faire ou avec valeur absolue de $DRR - DDL_{\text{minimal}}$ et $DROR + DCOR \leq DRR$ si copie à faire;

(56) - Inscription de l'objet courant **OR** comme prochain objet à réorganiser **POR** au planning de la région identifiée en valorisant le pointeur **POR** correspondant à l'adresse de l'objet **OR**, en augmentant la durée de temps consommé **DCRR** dans la région du temps de la réorganisation **DROR** et recalculant la durée gaspillée dans la région de réorganisation **DGRR** du fait des copies d'objets **OR** en fin de réorganisation ;

60) - Lancement, en l'absence de tout événement de remise à zéro **RAZR=1** survenu entre temps, dès qu'une région de traitement de réorganisation **RTR** identifiée est

disponible, de la réorganisation de l'objet **POR** ; Remise à zéro et nouvelle exécution du processus **IDPOR** ; et

- le processus **IDPOC** comportant les opérations suivantes :

5 (51) - Etablissement d'un planning rapide de réorganisation **PRR*** à l'aide de la liste **PRIOREORG** et du temps opérationnel de réorganisation résiduel disponible **TRR** et d'une liste de sélection **SELECT/OR** d'objets à réorganiser **OR** ordonnés par priorités décroissantes optimisées en fonction du gain de la réorganisation de chaque objet **OR**, ledit gain étant défini comme le produit d'un facteur représentatif du taux de désorganisation d'un objet **OR** par le temps de réorganisation de cet objet;

10 (52) - Etablissement d'un planning rapide de copie **PRC** à partir de la liste **SELECT/OR**, puis de la liste **PRIOCOPIE** et dans la limite du temps opérationnel de copie résiduel disponible **TRC**, d'une liste de sélection **SELECT/OC** en mémoire de type LIFO ou pile d'objets à copier **OC** empilés par priorités décroissantes;

15 (53') - Etablissement d'un rétroplanning de copie avec détermination des durées de copie disponibles en région copie, dépilage de la liste **SELECT/OC** de haut en bas et recherche d'une région copie par adéquation entre la durée copie disponible et la durée copie de l'objet courant à copier **OC** de façon à copier les objets les plus grands possibles et marquage de l'origine de réorganisation ou non de l'objet à copier ; suivi de

20 - l'identification de la région de traitement **RTC** de l'objet à copier par adéquation entre la durée de traitement disponible dans ladite région avec la durée de traitement nécessaire pour la copie de l'objet courant **OC**; suivie de

- l'inscription de l'objet courant **OC** comme le prochain objet à copier **POC** au planning de la région **RTC** identifiée en valorisant le pointeur **POC** correspondant à l'adresse de l'objet **OC** et en diminuant la durée de traitement disponible du temps de la copie de l'objet **OC** retenu comme **POC** ;

(60')- Lancement, en l'absence de tout évènement de remise à zéro **RAZC=1** survenu entre temps, dès qu'une région de traitement de copie **TRC** identifiée est disponible de la copie de l'objet **POC**; Remise à zéro et nouvelle exécution du processus **IDPOC**.

30 5. Méthode de gestion selon la revendication 4, caractérisée en ce qu'elle intègre dans les phases finales de processus **IDPOR** d'un objet **OR** une phase de recherche d'un ou plusieurs objets à réorganiser **OR** sans copie **OR** dans la liste **SELECT/OR** ou à défaut dans la liste **PRIOREORG** susceptibles d'être réorganisés dans la région de traitement **RTR** correspondante pendant l'intervalle du temps d'attente de la copie de l'objet **OR** après réorganisation.

35 6. Méthode de gestion selon l'une des revendications 4 et 5, caractérisée en ce que ladite phase (51) d'établissement d'un planning rapide de réorganisation **PRR*** comporte les opérations suivantes :

(511) - Initialisation : dans laquelle sont notamment annulés les indicateurs de sélection des objets **OR**, initialisé le compteur-limiteur de remises en causes, les temps résiduels de copie **TRC** et de réorganisation **TRR** et la durée minimale **E** de réorganisation et copie d'un objet **OR** ;

5 (512) - Boucle sur objets : dans laquelle notamment pour tout objet réel **OR** de la liste **PRIOREORG** pris dans l'ordre décroissant jusqu'à ce que $TRR < E$, une marque de 'sélection possible' de l'objet est posée;

(513)- Elimination des intersections d'objets : dans laquelle notamment sont éliminés de la 'sélection possible' les objets réels **OR** parties d'espaces de tables également
10 appelés objets globaux déjà sélectionnés en tout ou partie;

(514) - Remise en cause des choix précédents : dans laquelle notamment sont éliminés de la 'sélection possible' les objets **OR** dont le temps total de réorganisation et copie est supérieur au temps de réorganisation résiduel **TRR** et à défaut dans laquelle l'objet courant **OR** et les objets **OR** déjà sélectionnés dans la liste **SELECT/OR** sont soumis à
15 un processus d'optimisation en fonction du gain de réorganisation avec remplacement d'un objet déjà sélectionné par l'objet en cours avec essais, dans la limite du terme du compteur-limiteur, de combinaisons successives tendant à retenir pour être sélectionné au final dans **SELECT/OR** des objets de plus grand gain de réorganisation possible et à les associer aux autres objets **OR** de la 'sélection possible' de façon à ce
20 que le temps total des réorganisations de tous les objets **OR** retenus soit inférieur mais le plus proche possible du temps résiduel de réorganisation **TRR**.

(515) - Vérification de la suffisance du temps copie disponible dans le temps résiduel de copie **TRC** pour chaque objet **OR** en passe d'être finalement sélectionné pour la réorganisation compte tenu de la priorité de réorganisation de l'objet **OR** par rapport
25 aux objets à copier en préalable du fait de leurs priorités copie supérieures.

(516) - Sélection de l'objet courant **OR** dans la liste **SELECT/OR** : dans laquelle notamment sont positionnés les indicateurs de l'objet courant **OR** et de l'objet global associé, annulés éventuellement les indicateurs du ou des objets **OR** désélectionnés et diminuer **TRR** et **TRC** des temps respectifs de réorganisation et de copie.

30 7. Méthode de gestion selon l'une des revendications 4 à 6 caractérisée en ce que ladite phase (52) d'établissement d'un planning rapide de copie **PRC** comporte les opérations suivantes:

(521) - Initialisation : au sortir d'un planning de réorganisation **PRR*** et dans laquelle sont notamment vidées les piles de sélection de copie **PSC1** et **PSC2**, initialisés le
35 compteur-limiteur de remises en causes, le temps résiduel de copie **TRC** et la durée minimale **EC** de copie d'un objet à copier **OC** ;

(522) – Inspection des régions de réorganisation : Pour toutes les régions,

si région active, et si image copie de l'objet en cours de réorganisation à faire et/ou si image copie de tout objet **OR** de la liste **SELECT/OR** pris dans l'ordre des priorités décroissantes à faire: pour tout sous-objet susceptible d'être copié séparément, notamment une partition, soustraction du temps de copie du sous-objet de **TRC** et placement du sous-objet en haut de la première pile **PSC1** ;

5 (523) – traitement de la liste des objets à copier **PRIOCOPIE** : dans laquelle notamment pour tout objet à copier **OC** de la liste **PRIOCOPIE** pris dans l'ordre décroissant jusqu'à ce que **TRC < EC** et absent de **PSC1**, soustraction du temps de copie du sous-objet **OC** de **TRC** et placement du sous-objet **OC** en haut de la seconde pile **PSC2** ,

10 (524) – Empilage des piles de sélection : dans laquelle la pile **PSC1** est superposée à la pile **PSC2** pour constituer la liste **SELECT/OC**.

8. Méthode de gestion selon la revendication 7 caractérisée en ce que ladite phase (53) d'établissement d'un rétroplanning de copie dans le processus **IDPOR** comporte les opérations suivantes:

15 (531) Initialisation des régions de copie : dans laquelle au sortir d'un planning de réorganisation pour toutes les régions de copie, la durée consommée de la région **DCRC** est mise à zéro,

(532) Détermination du temps de copie à réserver **TCR** : dans laquelle pour tout objet réel de la liste **SELECT/OR** parcourue dans l'ordre des priorités croissantes pris comme objet courant à réorganiser **OR**,

20 -le temps à réserver pour copier **TCR** l'objet réel courant **OR** est mis à zéro ;
 -Si l'image copie de tout objet **OR** de la liste **SELECT/OR** pris dans l'ordre des priorités croissantes doit être faite: pour tout sous-objet susceptible d'être copié séparément, notamment partition, on recherche une région copie dont la durée consommée **DCRC** est minimale, on ajoute le temps de copie du sous-objet à la durée consommée de la région courante **DCRC** et on valorise le temps à réserver **TCR** pour copie de l'objet courant = **maximum** entre le temps à réserver **TCR** pour copie de l'objet courant et durée consommée dans la région **DCRC**.

30 9. Méthode de gestion selon la revendication 8, caractérisée en ce que ladite phase de rétroplanning de réorganisation (54) dans le processus **IDPOR** comporte les opérations suivantes :

(541) – Initialisation : dans laquelle pour toute région de réorganisation,
 - Durée de la région **DRR** = Durée de la fenêtre 'batch' **DFB** moins heure relative estimée de fin du traitement de réorganisation en cours dans la région **HRFR**

35 - Durée consommée dans la région **DCRR** = 0

- Durée gaspillée dans la région **DGRR** = 0

- Pointeur premier objet à réorganiser de la région courante = 0

(542) – Détermination du dernier délai de la région de réorganisation **DDRR** : dans laquelle pour tout objet réel **OR** sélectionné de la liste **SELECT/OR** dans l'ordre des priorités croissantes :

- 5 Pour chaque région si l'image copie de l'objet courant **OR** à faire, calcul du dernier délai de la région **DDRR** = durée de réorganisation de l'objet courant **DROR** + **maximum** entre le temps à réserver pour copie de l'objet courant **TCR** et le total de la durée consommée de la région **DCRR** + la durée gaspillée de la région **DGRR** ;
- Sinon, calcul du dernier délai de la région **DDR** = durée consommée de la région
- 10 **DCRR** + **maximum** entre la durée de réorganisation de l'objet courant **DROR** et la durée gaspillée de la région **DGRR**.

10. Méthode de gestion selon la revendication 9, caractérisée en ce que ladite phase d'identification de la région de réorganisation **RTR** (55) comporte les opérations suivantes : Selon le type d'objet courant à réorganiser **OR**

- 15 (551) – Recherche de la région de traitement de réorganisation susceptible d'accueillir de façon optimale un objet à réorganiser **OR** sans copie : comportant
- la recherche d'une région dont le dernier délai **DDRR** est **minimal**, et si dernier délai de la région courante **DDR** est **supérieur** à la durée de la région courante **DDR**, la recherche d'une région pour laquelle : la valeur de la durée de la région courante **DDR**
- 20 **moins** le dernier délai de la région courante **DDRR** est **minimale et positive ou nulle** ; sinon
- (552) – la recherche de la région de réorganisation susceptible d'accueillir de façon optimale un objet à réorganiser **OR** avec copie : comportant
- Recherche d'une région telle que : la valeur de la durée de la région **DDR** **moins**
- 25 dernier délai de la région courante **DDRR** est **minimale en valeur absolue**, et la valeur de la durée de réorganisation **DROR** **plus** le temps de copie objet courant **TCR** est inférieur ou égal à la durée de la région **DDR**.

11. Méthode de gestion selon la revendication 10, caractérisée en ce que ladite phase d'inscription de l'objet **OR** au planning de la région de réorganisation **RTR** (56) comporte les opérations suivantes :

- 30 (561) – Inscription de l'objet **OR** comme prochain objet à réorganiser **POR** : dans laquelle si l'image copie de l'objet courant à faire ou durée gaspillée de la région courante égale zéro, on valorise le **pointeur premier objet à réorganiser POR de la région courante à l'adresse de l'objet courant OR**,
- (562) - mise à jour **DCRR** et **DGRR** : dans laquelle on ajoute la durée de réorganisation
- 35 de l'objet courant **DROC** à la durée consommée de la région courante **DCRR**, et on valorise la durée gaspillée de la région courante **DGRR** à la valeur du dernier délai de la région courante **DDRR** **moins** la durée consommée de la région courante **DCRR**.

12. Méthode de gestion selon l'une des revendications 4 à 7 caractérisée en ce que ladite phase (53') d'établissement d'un rétroplanning de copie dans le processus **IDPOC** comporte les opérations suivantes:

(531') Initialisation des régions de copie : dans laquelle pour toutes les régions,

- 5 - la durée de région **DRCOP** est valorisée à la durée fenêtre « batch » **DFB** moins l'heure relative estimée de fin de traitement copie en cours **HRFC**,
- la durée consommée de la région **DCRCOP** est mise à zéro,
- le pointeur du premier objet à copier **POC** de la région courante est mis à zéro

(532') Dépilage et identification du prochain objet à copier **POC** : Dans la laquelle,

- 10 Pour tout objet de la pile **PSC1/PSC2** parcourue de haut en bas un objet courant à copier **OC** est extrait lorsque le temps de copie de l'objet courant **TCOC** est inférieur ou égal à la longue **DRCOP**;

On cherche pour identifier la région de traitement copie une région dont la durée consommée **DCRCOP** est minimale ; dans l'affirmative si le temps de copie de l'objet

- 15 courant **TCOC** est supérieur à **DRCOP** moins **DCRCOP** on cherche une région pour laquelle **DRCOP** moins **DCRCOP** moins **TCOC** est minimale, positive ou nulle,

Si aucune région ne convient, on choisit une région à copier pour laquelle **DRCOP** moins **DCRCOP** moins **TCOC** est négative et minimale en valeur absolue et pour laquelle **TCOC** est inférieur ou égal à **DRCOP** ;

- 20 Si l'objet courant appartient à la seconde pile **PSC2**, soit un objet à copier non issu d'une réorganisation, le pointeur de l'objet **POC** pour la région courante choisie est valorisé à l'adresse de l'objet courant **OC** de façon à inscrire l'objet **POC** au planning de la région de traitement copie **RTC** identifiée ; et

- 25 Pour terminer, on ajoute la durée copie **TCOC** à la durée consommée de la région courante **DCRCOP**.

- 30 13. Méthode de gestion selon l'une de revendications précédentes, caractérisée en ce que lesdits événements internes au système **RAZR=1** et **RARC=1** entraînant la remise en cause de la sélection des objets **POR** ou **POC** sont notamment constitués respectivement par la fin d'une tâche de réorganisation dans une région de traitement de réorganisation pour les objets **POR** ou par la fin d'une tâche de copie dans une région de traitement de copie pour les objets **POC**, une copie prioritaire en fin de réorganisation et/ou la libération d'une zone de traitement pour les objets **POR** et **POC**.

- 35 14. Méthode de gestion des réorganisations d'un ensemble de bases de données indexées d'un système informatique d'information adaptée à la réorganisation « en ligne » et destinée à travailler en association avec une méthode de gestion selon l'une des revendications précédentes, caractérisée en ce qu'elle comporte au moins les phases opérationnelles suivantes:

- une phase d'analyse instantanée d'au moins un objet à réorganiser **OR** parmi les objets susceptibles de réorganisation, notamment bases de données, partitions et/ou index à réorganiser, et d'estimation du surcoût associé au niveau de désorganisation de l'objet à réorganiser **OR**;
- 5 · une phase d'estimation instantanée du coût de la réorganisation en ligne en fonction de la taille dudit objet à réorganiser **OR** et du taux de désorganisation dudit objet à réorganiser **OR** ; et
- une phase de détermination du seuil **Ds** de désorganisation minimal de l'objet à réorganiser **OR** au dessus duquel seuil la réorganisation « en ligne » peut être lancée
- 10 pour cet objet **OR** et le cas échéant du lancement effectif de cette réorganisation en ligne, le seuil **Ds** correspondant sensiblement au minimum du coût total du surcoût estimé de la désorganisation et du coût estimé de la réorganisation pour l'objet **OR** considéré.

15 15. Méthode de gestion des réorganisations « en ligne » selon la revendication 14, caractérisée en ce que le lancement de la réorganisation « en ligne » est retardé pour attendre une fenêtre temporelle d'activité réduite de la base de données concernée.

16. Méthode de gestion des réorganisations « en ligne » et « hors ligne » selon l'une des revendications 14 et 15, caractérisée en ce que pour un objet à réorganiser **OR** sélectionné comme premier objet à réorganiser **POR** la priorité est donnée pour la

20 réorganisation « hors ligne », l'intervention de la réorganisation « en ligne » n'étant requise qu'après dépassement du seuil **Ds** pour l'objet **POR** considéré.

17. Méthode de gestion des réorganisations « en ligne » selon l'une des revendications 14 à 16 et utilisée pour la réorganisation des bases de données indexées de type espace de tables, la méthode étant caractérisée en ce que pour un

25 objet à réorganiser **OR**, le seuil **Ds** est défini par une approximation donnée par la formule **F5'** suivante :

$$R = Ds^2 r Co^* / 2.l.c = 1$$

lorsque que **U** moyenne horaire du nombre de mises à jour sur l'objet **OR** est faible, et dans laquelle :

30 **r** désigne le nombre moyen de lignes ou clé-RID par page de l'objet entier,

l désigne le nombre d'insertions de lignes, ou de clé-RID pour les index, à l'heure dans l'objet **OR**,

Co* désigne le surcoût horaire lorsque l'objet est totalement désorganisé

35 **c** désigne un paramètre machine, le rapport du temps d'accès **E/S** par la taille de l'objet **Nbp**.

18. Méthode de gestion des réorganisations « en ligne » selon la revendication 17 et utilisée pour la réorganisation des index, caractérisé en ce que pour **un index**, le seuil **Dsl** est défini par une approximation donnée par la formule **F7'** suivante :

$$R = Dsl^2 r (t^*+1) Nbp.V / [2(V+3)] = 1$$

- 5 lorsque que **U** moyenne horaire du nombre de mises à jour sur l'index est faible, et dans laquelle :

Ts désigne le temps de déplacement de bras, **TI** le temps de latence (1/2 tour du bras).

c = (Ts+ 3.TI)/Nbp où **Ts** en **E/S** désigne le temps de déplacement du bras du disque, et **TI** le temps de latence, 1/2 tour du bras, et **V = Ts/TI**.

- 10 **t*** désigne le taux de relecture de l'objet, c'est à dire le nombre moyen de fois qu'une entrée sera relue.

19. Méthode de gestion des réorganisations « en ligne » selon la revendication 17 et utilisée pour la réorganisation des bases de données indexées de type espace de tables monotable à **n** index, caractérisé en ce que le seuil **DsT** est défini pour **un**

15 **espace monotable** par une approximation donnée par la formule **F13'** suivante :

$$R = r.Nbp[DsT^2.L (V+1) + V. S (DsT^2-bi^2) (t^*i+1)] / (2V+6) = 1$$

lorsque que **U**, moyenne horaire du nombre de mises à jour sur l'espace monotable est faible, et dans laquelle :

- 20 **L** désigne la proportion de lignes accédées par parcours séquentiel sur le nombre de lignes créées,

S représente le signe mathématique Sigma, somme des expressions **f(xi)** pour **i** index allant de 1 à **n**,

bi représentent les valeurs instantanées de **D**, taux de désorganisation de l'espace monotable, lors des dernières réorganisations des **n** index, et

- 25 **t*i** représentent les taux d'accès directs au travers de chaque index **i**.

20. Méthode de gestion des réorganisations dans un ensemble de bases de données indexées d'un système informatique d'information selon l'une des revendications précédentes, caractérisée en ce qu'elle est appliquée à la réorganisation de bases de données indexées de type espace de tables.

- 30 21. Système informatique de type système d'information caractérisé en ce qu'il comporte un ensemble de bases de données indexées et les moyens matériels et logiciels adaptés pour la mise en œuvre de la méthode de gestion des réorganisations de base de données indexées selon l'une des revendications précédentes, notamment de bases de données de type espace de tables et/ou d'index.

FIG.1

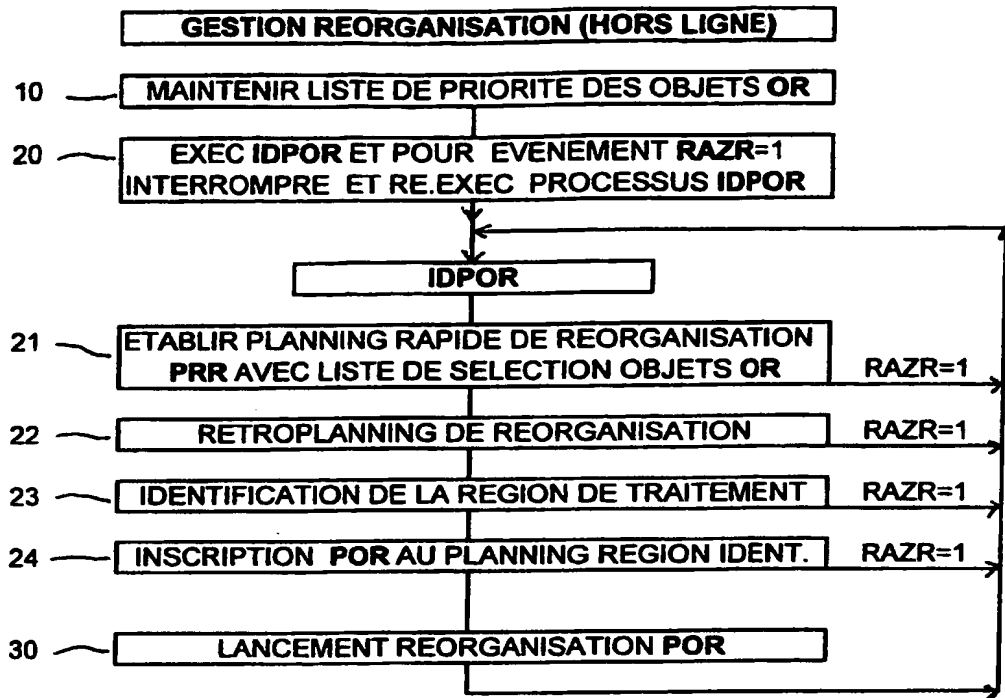


FIG.2

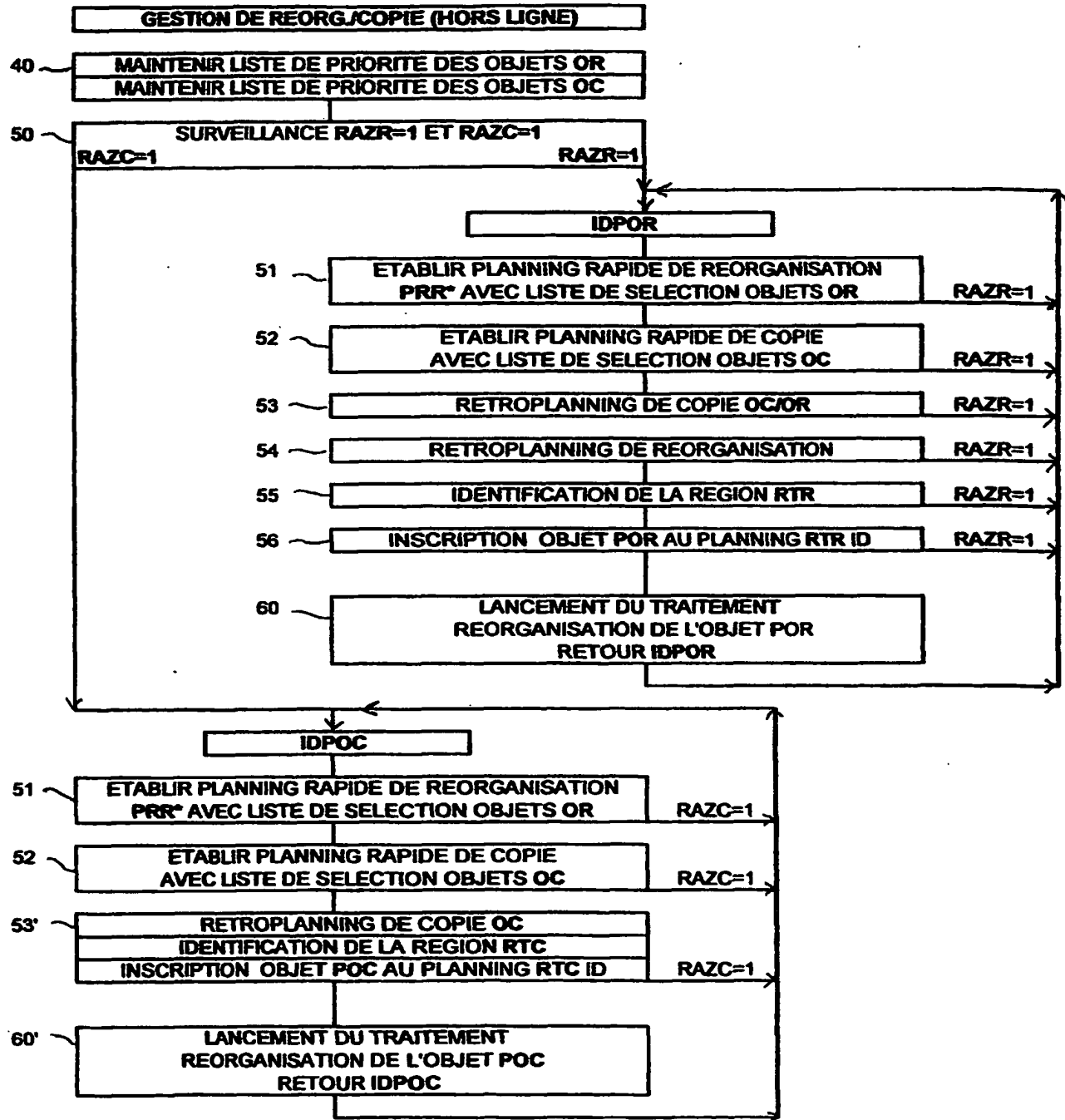


FIG.3

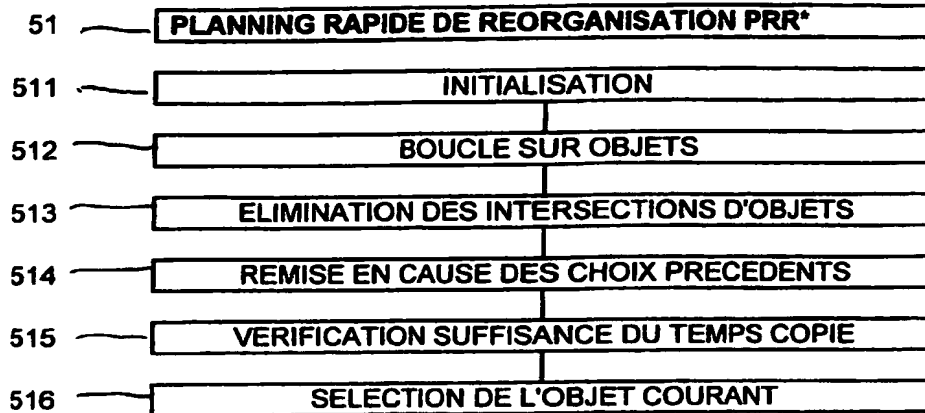
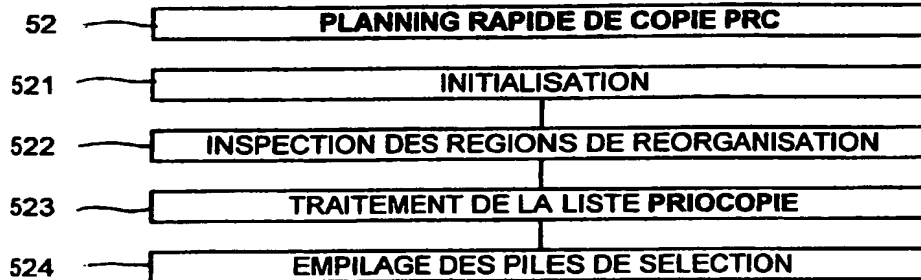
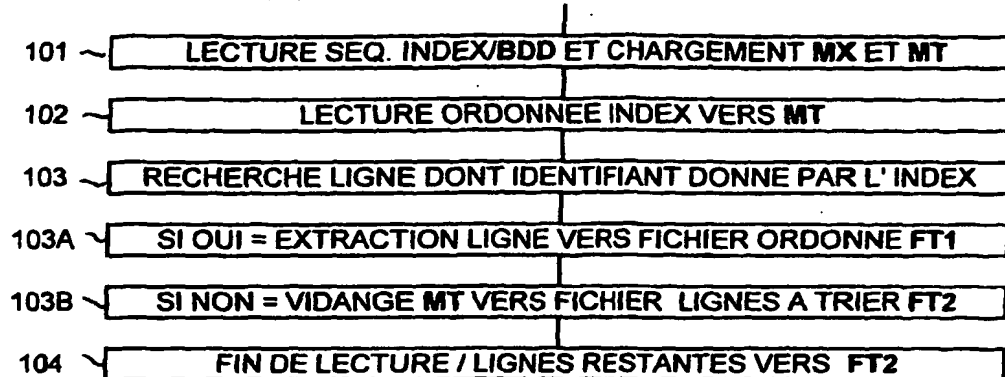


FIG.4



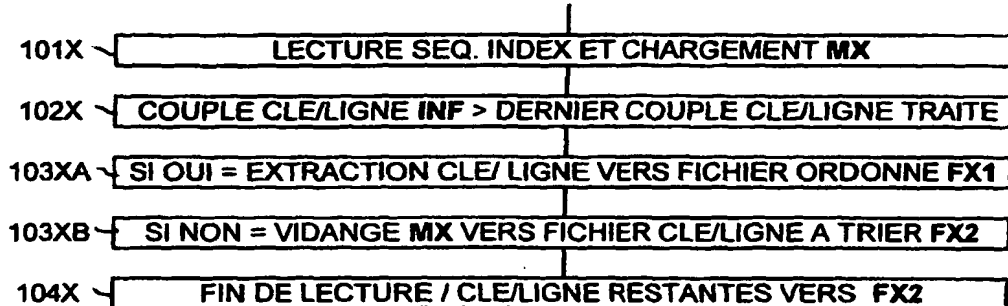
LEC/BD/REORG/BD
LECTURE DE BASE DE DONNEES INDEXEE STRUCTUREE
EN LIGNES - REORGANISATION BASE DE DONNEES

FIG.5A



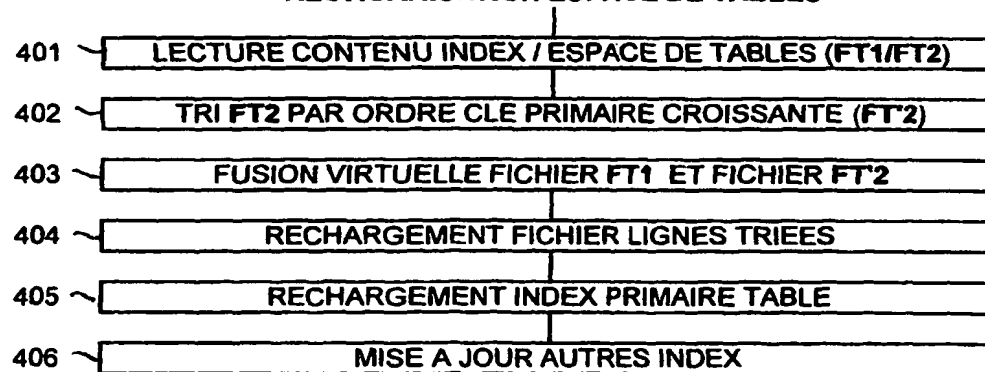
LEC/IND/REORG/IND
LECTURE D'INDEX DE BASE DE DONNEES INDEXEE
STRUCTUREE EN LIGNES - REORGANISATION INDEX

FIG.5B



REORG/TABLE
REORGANISATION ESPACE DE TABLES

FIG. 8A



REORG/IND(ET)
REORGANISATION INDEX

FIG. 8B

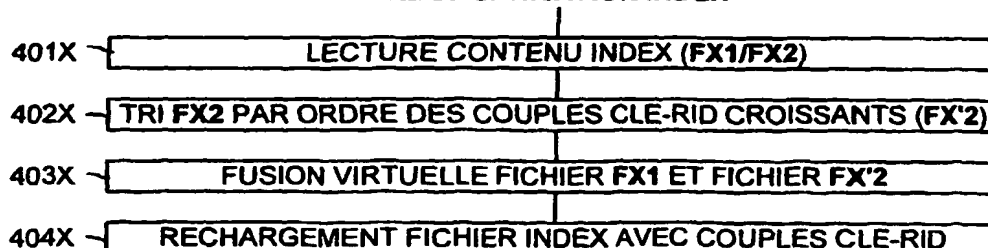


FIG. 6A

LEC/IND/REORG/TABLE
DECHARGEMENT INDEX - REORGANISATION ESP. DE TABLES

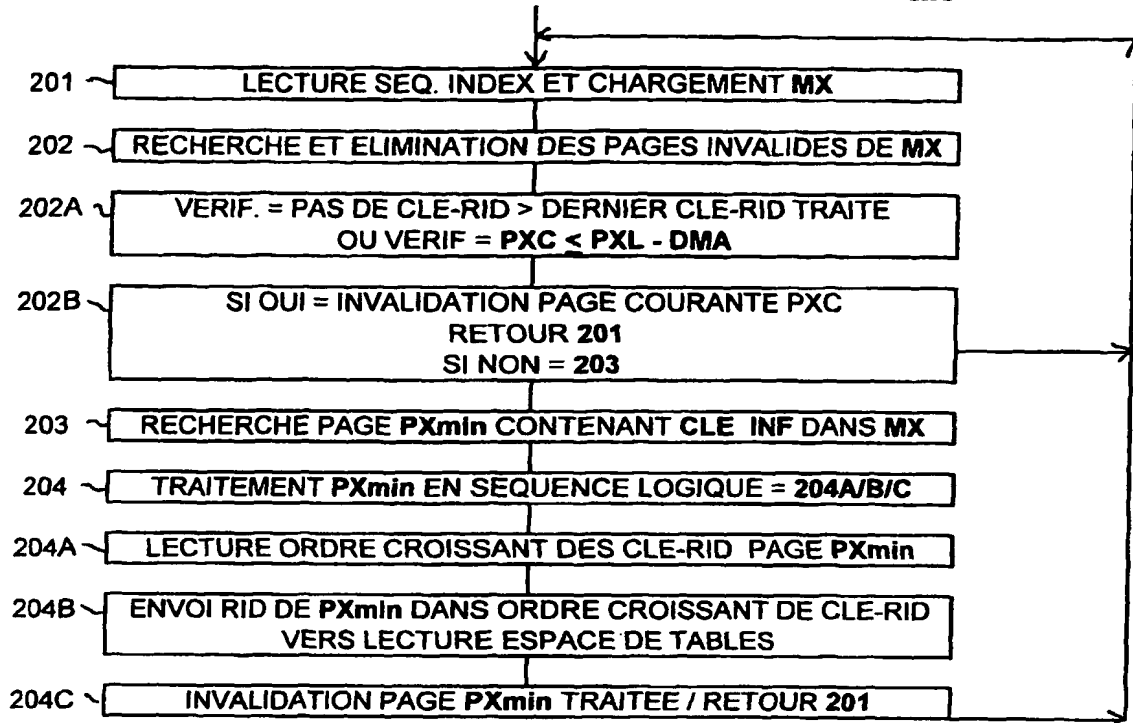
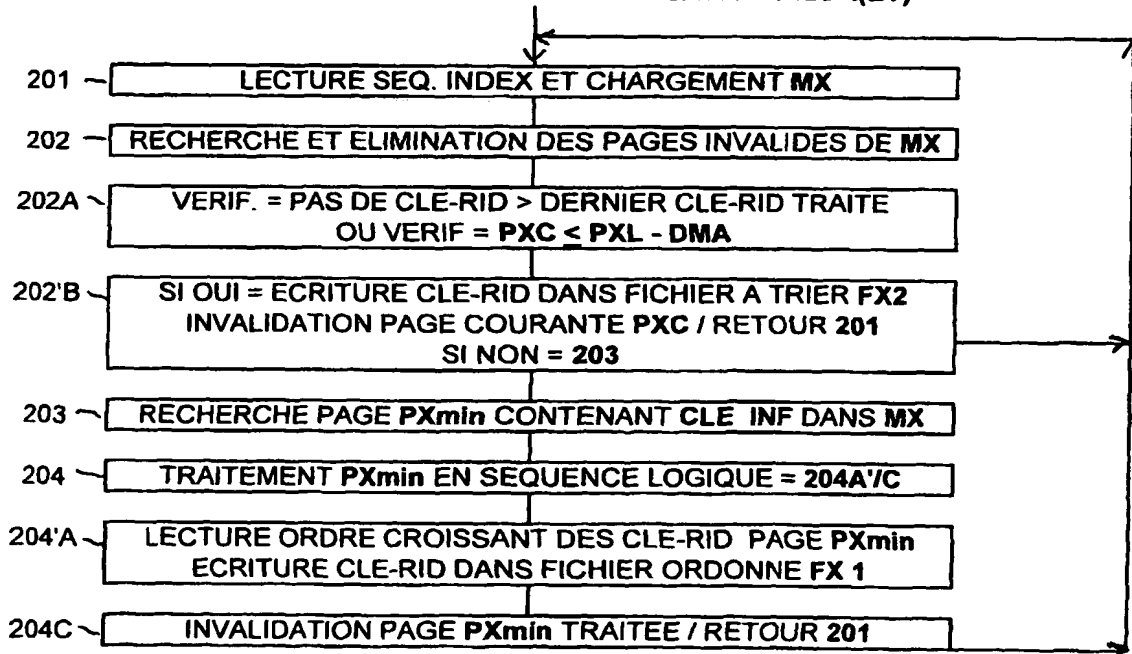


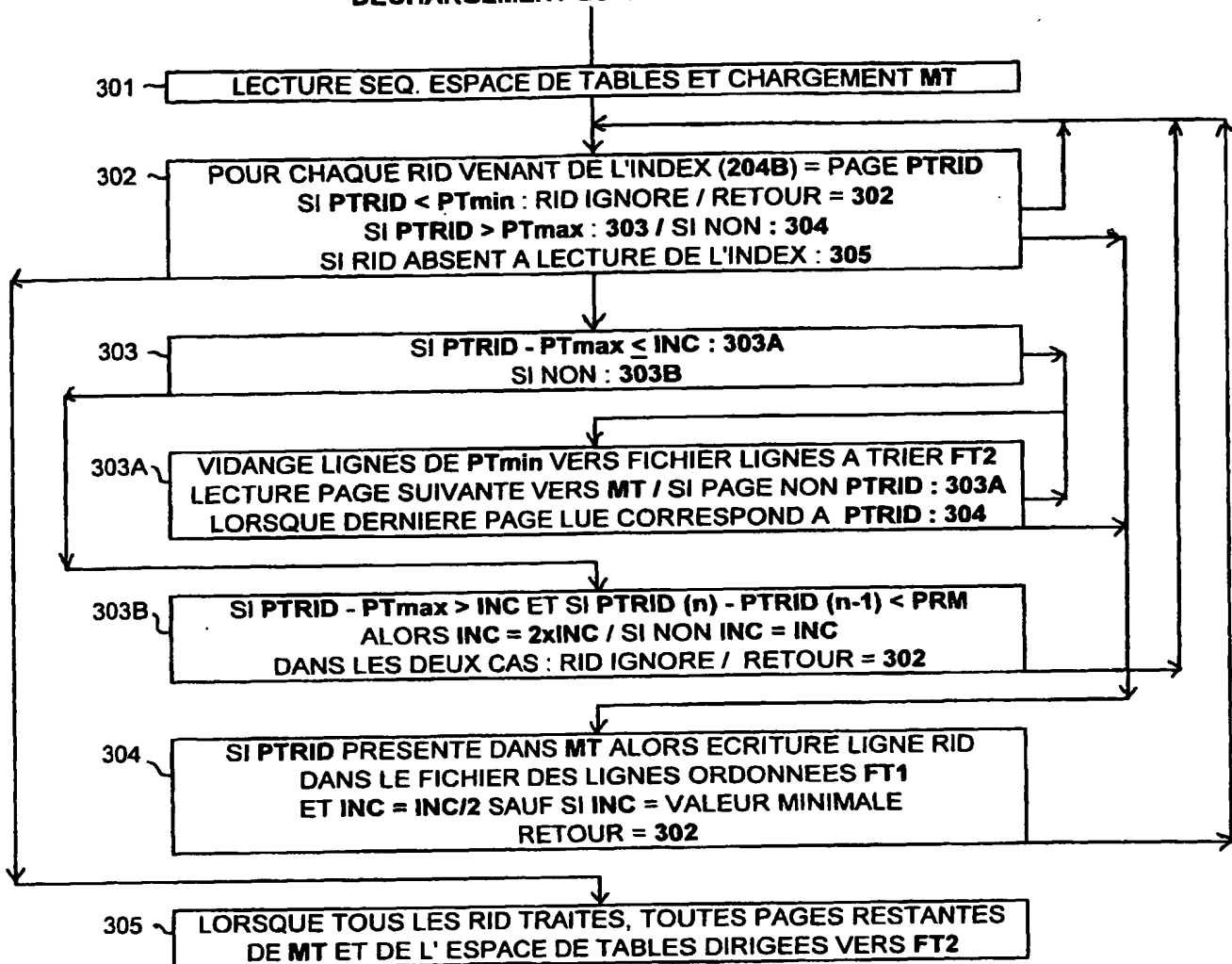
FIG. 6B

LEC/IND/REORG/IND(ET)
DECHARGEMENT INDEX - REORGANISATION INDEX(ET)



LEC/TABLE/REORG/TABLE

DECHARGEMENT ESPACE DE TABLES



INTERNATIONAL SEARCH REPORT

International Application No

PCT/FR 03/01076

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 881 379 A (BEIER HARLEY AL ET AL) 9 March 1999 (1999-03-09) column 3, line 40 -column 4, line 10 ---	1-21
A	SOCKUT G H ET AL: "A METHOD FOR ON-LINE REORGANIZATION OF A DATABASE" IBM SYSTEMS JOURNAL, IBM CORP. ARMONK, NEW YORK, US, vol. 36, no. 3, 1997, pages 411-436, XP000701880 ISSN: 0018-8670 figure 4 page 417 -page 418 page 430, right-hand column ---	1-21
A	US 5 887 274 A (AL-EISA EISA A ET AL) 23 March 1999 (1999-03-23) figures 1A,1B --- -/--	1-21



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *&* document member of the same patent family

Date of the actual completion of the international search

13 August 2003

Date of mailing of the international search report

26/08/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl
Fax: (+31-70) 340-3016

Authorized officer

Michalski, S

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/FR 03/01076

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MARUYAMA K ET AL: "Optimal reorganization of distributed space disk files" COMMUNICATIONS OF THE ACM, NOV. 1976, USA, vol. 19, no. 11, pages 634-642, XP002236355 ISSN: 0001-0782 page 636, right-hand column -----	1-21
A	BARU C ET AL: "Data reorganization in parallel database systems" ADVANCES IN PARALLEL AND DISTRIBUTED SYSTEMS, 1993., PROCEEDINGS OF THE IEEE WORKSHOP ON PRINCETON, NJ, USA 6 OCT. 1993, LOS ALAMITOS, CA, USA, IEEE COMPUT. SOC, 1993, pages 102-107, XP010223428 ISBN: 0-8186-5250-0 page 102, right-hand column -page 103, left-hand column -----	1-21

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 03/01076

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5881379	A	09-03-1999	NONE	
US 5887274	A	23-03-1999	US 5758357 A	26-05-1998
			US 5517641 A	14-05-1996
			US 5408654 A	18-04-1995

RAPPORT DE RECHERCHE INTERNATIONALE

Dén. de l'Invent. Nationale No
PCT/FR 03/01076

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
CIB 7 G06F17/30

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 7 G06F

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	US 5 881 379 A (BEIER HARLEY AL ET AL) 9 mars 1999 (1999-03-09) colonne 3, ligne 40 -colonne 4, ligne 10 ---	1-21
A	SOCKUT G H ET AL: "A METHOD FOR ON-LINE REORGANIZATION OF A DATABASE" IBM SYSTEMS JOURNAL, IBM CORP. ARMONK, NEW YORK, US, vol. 36, no. 3, 1997, pages 411-436, XP000701880 ISSN: 0018-8670 figure 4 page 417 -page 418 page 430, colonne de droite ---	1-21
A	US 5 887 274 A (AL-EISA EISA A ET AL) 23 mars 1999 (1999-03-23) figures 1A,1B ---	1-21
-/-		

☒ Voir la suite du cadre C pour la fin de la liste des documents

☒ Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- *A* document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- *E* document antérieur, mais publié à la date de dépôt international ou après cette date
- *L* document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- *O* document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- *P* document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

- *T* document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention
- *X* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément
- *Y* document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier
- *A* document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

13 août 2003

Date d'expédition du présent rapport de recherche internationale

26/08/2003

Nom et adresse postale de l'administration chargée de la recherche internationale
Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Michalski, S

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale No

PCT/FR 03/01076

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	MARUYAMA K ET AL: "Optimal reorganization of distributed space disk files" COMMUNICATIONS OF THE ACM, NOV. 1976, USA, vol. 19, no. 11, pages 634-642, XP002236355 ISSN: 0001-0782 page 636, colonne de droite ---	1-21
A	BARU C ET AL: "Data reorganization in parallel database systems" ADVANCES IN PARALLEL AND DISTRIBUTED SYSTEMS, 1993., PROCEEDINGS OF THE IEEE WORKSHOP ON PRINCETON, NJ, USA 6 OCT. 1993, LOS ALAMITOS, CA, USA, IEEE COMPUT. SOC, 1993, pages 102-107, XP010223428 ISBN: 0-8186-5250-0 page 102, colonne de droite -page 103, colonne de gauche -----	1-21

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale No

PCT/FR 03/01076

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
US 5881379	A	09-03-1999	AUCUN	
US 5887274	A	23-03-1999	US 5758357 A	26-05-1998
			US 5517641 A	14-05-1996
			US 5408654 A	18-04-1995